11-30-00 09701674 PCT

526 Rec'd PCT/PTO 28 NOV 2000

FORM PTO-1390 U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371		ATTORNEY'S DOCKET NUMBER **PF-0539 USN** U.S. APPLICATION NO. (If known, see 37 CFR 1.5) **TO BE ASSIGNED** 09/701674
INTERNATIONAL APPLICATION NO. **PCT/US99/13281**	INTERNATIONAL FILING DATE **11 June 1999**	PRIORITY DATE CLAIMED **12 June 1998**

TITLE OF INVENTION
CELL CYCLE REGULATION PROTEINS

APPLICANT(S) FOR DO/EO/US
INCYTE PHARMACEUTICALS, INC.; LAL, Preeti; YUE, Henry; TANG, Y. Tom; HILLMAN, Jennifer L.; BANDMAN, Olga; CORLEY, Neil C.; GUEGLER, Karl J.; GORGONE, Gina A.; BAUGHN, Mariah R.; PATTERSON, Chandra; LU, Dyung Aina M.

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. ☒ This is the **FIRST** submission of items concerning a filing under 35 U.S.C. 371.
2. ☐ This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S.C. 371.
3. ☐ This is an express request to promptly begin national examination procedures (35 U.S.C. 371 (f)).
4. ☐ The US has been elected by the expiration of 19 months from the priority date (PCT Article 31).
5. ☒ A copy of the International Application as filed (35 U.S.C. 371(c)(2))
   a. ☐ is attached hereto (required only if not communicated by the International Bureau)
   b. ☐ has been communicated by the International Bureau.
   c. ☒ is not required, as the application was filed in the United States Receiving Office (RO/US).
6. ☐ An English language translation of the International Application as filed (35 U.S.C. 371(c)(2)).
7. ☒ Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3))
   a. ☐ are attached hereto (required only if not communicated by the International Bureau).
   b. ☐ have been communicated by the International Bureau.
   c. ☐ have not been made; however, the time limit for making such amendments has NOT expired.
   d. ☒ have not been made and will not be made.
8. ☐ An English language translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).
9. ☐ An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).
10. ☐ An English language translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).

Items 11 to 16 below concern document(s) or information included:

11. ☐ An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
12. ☐ An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.27 and 3.31 is included.
13. ☐ A FIRST preliminary amendment.
    ☐ A SECOND or SUBSEQUENT preliminary amendment.
14. ☐ A substitute specification.
15. ☐ A change of power of attorney and/or address letter.
16. ☒ Other items or information:

1) Transmittal Letter (2 pp, in duplicate)
2) Return Postcard
3) Express Mail Label No.: **EL 579 976 120 US**

U.S. APPLICATION NO. (if known, see 37 CFR 1.5) NEW APPLICATION	INTERNATIONAL APPLICATION NO.: PCT/US99/13281	ATTORNEY'S DOCKET NUMBER PF-0539 USN

09/701674

17. ☐ The following fees are submitted:

BASIC NATIONAL FEE (37 CFR 1.492(a)(1)-(5):

Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO and International Search Report not prepared by the EPO or JPO.....**$1000.00**

International preliminary examination fee (37 CFR 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO..**$860.00**

International preliminary examination fee (37 CFR 1.482) not paid to USPTO but international search fee (37 CFR 1.445(a)(2)) paid to USPTO.....**$710.00**

☒International preliminary examination fee paid to USPTO (37 CFR 1.482) but all claims did not satisfy provisions of PCT Article 33(1)-(4).....**$690.00**

International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(1)-(4).....**$100.00**

ENTER APPROPRIATE BASIC FEE AMOUNT =				$690.00	
Surcharge of **$130.00** for furnishing the oath or declaration later than ☐ 20 ☐ 30 months from the earliest claimed priority date (37 CFR 1.492(e)).				$	
CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE		
Total Claims	20 =	0	X **$ 18.00**	$	
Independent Claims	2 =	0	X **$ 80.00**	$	
MULTIPLE DEPENDENT CLAIM(S) (if applicable)			+ **$270.00**	$	
TOTAL OF ABOVE CALCULATIONS =				$690.00	
☐ Applicant claims small entity status. See 37 CFR 1.27. The fees indicated above are reduced by 1/2.				$	
SUBTOTAL =				$690.00	
Processing fee of **$130.00** for furnishing the English translation later than ☐ 20 ☐ 30 months from the earliest clailmed priority date (37 CFR 1492(f)). +				$	
TOTAL NATIONAL FEE =				$690.00	
Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by the appropriate cover sheet (37 CFR 3.28, 3.31). **$40.00** per property +				$	
TOTAL FEES ENCLOSED =				$690.00	
				Amount to be Refunded:	$
				Charged:	$

a. ☐ A check in the amount of $_____ to cover the above fees is enclosed.
b. ☒ Please charge my Deposit Account No. 09-0108 in the amount of $ 690.00 to cover the above fees.
c. ☒ The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 09-0108. A duplicate copy of this sheet is enclosed.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

INCYTE GENOMICS, INC.
3160 Porter Drive
Palo Alto, CA 94304

SIGNATURE

NAME: Diana Hamlet-Cox

REGISTRATION NUMBER: 33,302

DATE: 28 November 2000

PROTEINS REGULATING GENE EXPRESSION

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of proteins regulating gene expression and to the use of these sequences in the diagnosis, treatment, and prevention of reproductive disorders, nervous disorders, and diseases associated with cell proliferation and differentiation, including cancer, immune disorders, and developmental disorders.

BACKGROUND OF THE INVENTION

Multicellular organisms are comprised of diverse cell types that differ dramatically both in structure and function. The identity of a cell is determined by its characteristic pattern of gene expression, and different cell types express overlapping but distinctive sets of genes throughout development. Spatial and temporal regulation of gene expression is critical for the control of cell proliferation, cell differentiation, apoptosis, senescence, and other cellular processes that contribute to organismal development. Furthermore, gene expression is regulated in response to extracellular signals that mediate cell-cell communication and coordinate the activities of different cell types. Appropriate gene regulation also ensures that cells function efficiently by expressing only those genes whose functions are required at a given time.

Transcription Factors

Transcriptional regulatory proteins are essential for the control of gene expression. Some of these proteins function as transcription factors that initiate, activate, repress, or terminate gene transcription. Transcription factors generally bind to the promoter, enhancer, and upstream regulatory regions of a gene in a sequence-specific manner, although some factors bind regulatory elements within or downstream of the internal coding region of a gene. Transcription factors may bind to a specific region of DNA singly or as a complex with other accessory factors. (Reviewed in Lewin, B. (1990) Genes IV, Oxford University Press, New York, NY, pp. 554-570.)

The double helix structure and repeated sequences of DNA create topological and chemical features which can be recognized by transcription factors. These features are hydrogen bond donor and acceptor groups, hydrophobic patches, major and minor grooves, and regular, repeated stretches of sequence which induce distinct bends in the helix. Typically, transcription factors recognize specific DNA sequence motifs about 20 nucleotides in length or less. Multiple, adjacent transcription factor-binding motifs may be required for gene regulation.

Many transcription factors incorporate DNA-binding structural motifs which comprise either α helices or ß sheets that bind to the major groove of DNA. Four well-characterized

structural motifs are helix-turn-helix, zinc finger, leucine zipper, and helix-loop-helix. Proteins containing these motifs may act alone as monomers, or they may form homo- or heterodimers that interact with DNA.

The helix-turn-helix motif consists of two α helices connected at a fixed angle by a short chain of amino acids. One of the helices binds to the major groove. Helix-turn-helix motifs are exemplified by the homeobox motif which is present in homeodomain proteins. These proteins are critical for specifying the anterior-posterior body axis during development and are conserved throughout the animal kingdom. The Antennapedia and Ultrabithorax proteins of Drosophila melanogaster are prototypical homeodomain proteins (Pabo, C.O. and R.T. Sauer (1992) Ann. Rev. Biochem. 61:1053-1095).

The zinc finger motif, which binds zinc ions, generally contains tandem repeats of about 30 amino acids consisting of periodically spaced cysteine and histidine residues. Examples of this sequence pattern, designated C2H2 and C3HC4, have been described (Lewin, supra.). Zinc finger proteins each contain an α helix and an antiparallel ß sheet whose proximity and conformation are maintained by the zinc ion. Contact with DNA is made by the arginine preceding the α helix and by the second, third, and sixth residues of the a helix. Variants of the zinc finger motif include poorly defined cysteine-rich motifs which bind zinc or other metal ions. These motifs may not contain histidine residues and are generally nonrepetitive.

The bromodomain signature is an additional conserved region of about 70 amino acids found in a number of transcriptional regulatory proteins (ExPASy PROSITE document PS00633; Haynes, S.R. et al. (1992) Nucleic Acids Res. 20:2603). Although the exact function of this domain is unclear, it is found in the DNA-binding region of the thyroid hormone receptor coactivating protein. The thyroid hormone receptor is a member of the steroid/thyroid receptor superfamily that regulates the expression of many target genes through binding to thyroid hormone response elements (Tsuyoshi, M. et al (1997) J. Biol. Chem. 272:29834-29841). The bromodomain signature is also found in eukaroytic transcriptional initiation factor, TFIID, a protein essential for progression of the G1 phase of the cell cycle.

The leucine zipper motif comprises a stretch of amino acids rich in leucine which can form an amphipathic α helix. This structure provides the basis for dimerization of two leucine zipper proteins. The region adjacent to the leucine zipper is usually basic, and upon protein dimerization, is optimally positioned for binding to the major groove.

The helix-loop-helix motif consists of a short α helix connected by a loop to a longer α helix. The loop is flexible and allows the two helices to fold back against each other and to bind to DNA. The transcription factor Myc contains a prototypical HLH motif.

Most transcription factors contain characteristic DNA binding motifs including, but not limited to, those described above. Variations on the above motifs and new motifs have been and are currently being characterized (Faisst, S. and S. Meyer (1992) Nucl. Acids Res. 20:3-26).

Chromatin Associated Proteins

In the nucleus, DNA is packaged into chromatin, the compact organization of which limits the accessibility of DNA to transcription factors and plays a key role in gene regulation (Lewin, supra, pp. 409-410). The compact structure of chromatin is determined and influenced by chromatin-associated proteins such as the histones, the high mobility group (HMG) proteins, and the chromodomain proteins. There are five classes of histones, H1, H2A, H2B, H3, and H4, all of which are highly basic, low molecular weight proteins. The fundamental unit of chromatin, the nucleosome, consists of 200 base pairs of DNA associated with two copies each of H2A, H2B, H3, and H4. H1 links adjacent nucleosomes. HMG proteins are low molecular weight, non-histone proteins that may play a role in unwinding DNA and stabilizing single-stranded DNA. Chromodomain proteins play a key role in the formation of highly compacted heterochromatin, which is transcriptionally silent.

RNA-Associated Proteins

Much of the regulation of gene expression in eukaroytic cells occurs at the posttranscriptional level. Messenger RNAs (mRNA) which are produced in the cell nucleus from primary transcripts of protein-encoding genes are processed and transported to the cytoplasm where the protein synthesis machinery is located. RNA-associated proteins are a group of proteins that participate in the processing, splicing, editing, transport, localization, translation, stability, and posttranscriptional regulation of mRNAs. Such proteins include RNA helicases, splicing factors, nucleases, and translational regulatory proteins. In addition, the nucleolus is a highly organized subcompartment of the nucleus which contains protein machinery specifically dedicated to the transcription and processing of ribosomal RNAs. The RNA-binding activity of these proteins is mediated by a series of RNA-binding motifs identified within them. These domains include the RNP motif, the arginine-rich motif, the RGG box, and the KH motif. (Reviewed in Burd, C. G. and Dreyfuss, G. (1994) Science 265:615 - 621.) The RNP motif is the most widely found and best characterized of these motifs. It is composed of 90-100 amino acids which form an RNA-binding domain, and is found in one or more copies in proteins that bind pre-mRNA, mRNA, pre-ribosomal RNA, and small nuclear RNAs. The RNP motif is composed of two short sequences (RNP-1 and RNP-2) and a number of other mostly hydrophobic, conserved amino acids interspersed throughout the motif (Burd, supra; ExPASy PROSITE document PD0C0030).

Diseases and disorders related to gene regulation

Many neoplastic disorders in humans can be attributed to inappropriate gene expression. Malignant cell growth may result from either excessive expression of tumor promoting genes or insufficient expression of tumor suppressor genes (Cleary, M.L. (1992) Cancer Surv. 15:89-104). Chromosomal translocations may also produce chimeric loci which fuse the coding sequence of one gene with the regulatory regions of a second unrelated gene. Such an arrangement likely results in inappropriate gene transcription. The Wilms tumor suppressor gene product, WT1, is a protein containing a DNA-binding domain consisting of four zinc fingers and a proline-glutamine rich region capable of regulating transcription (ExPASy PROSITE document PR00049). Deletions of the WT1 gene, or point mutations which destroy the DNA-binding activity of the protein, are assiociated with development of the pediatric nephroblastoma, Wilms tumor, and Denys-Drash syndrome (Rauscher, F.J. (1993) FASEB J. 7:896-903).

Certain proteins enriched in glutamine are associated with various neurological disorders including spinocerebellar ataxia, bipolar effective disorder, schizophrenia and autism (Margolis, R.L. et al. (1997) Human Genetics 100:114-122). These proteins contain regions with as many as 15 or more consecutive glutamine residues and may function as transcription factors with a potential role in regulation of neurodevelopment or neuroplasticity.

The immune system responds to infection or trauma by activating a cascade of events that coordinate the progressive selection, amplification, and mobilization of cellular defense mechanisms. A complex and balanced program of gene activation and repression is involved in this process. However, hyperactivity of the immune system as a result of improper or insufficient regulation of gene expression may result in considerable tissue or organ damage. This damage is well documented in immunological responses associated with arthritis, allergens, heart attack, stroke, and infections (Harrison's Principles of Internal Medicine, 13/e, McGraw Hill, Inc. and Teton Data Systems Software, 1996). In particular, a zinc finger protein termed Staf50 (for Stimulated trans-acting factor of 50 kDa) is a transcriptional regulator and is induced in various cell lines by interferon-I and -II. Staf50 appears to mediate the antiviral activity of interferon by down-regulating the viral transcription directed by the long terminal repeat promoter region of human immunodeficiency virus type-1 in transfected cells (Tissot, C. (1995) J. Biol. Chem. 270:14891-14898).

Furthermore, the generation of multicellular organisms is based upon the induction and coordination of cell differentiation at the appropriate stages of development. Central to this process is differential gene expression, which confers the distinct identities of cells and tissues throughout the body. Failure to regulate gene expression during development could result in developmental disorders.

The discovery of new proteins regulating gene expression and the polynucleotides

encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of reproductive disorders, nervous disorders, and diseases associated with cell proliferation and differentiation, including cancer, immune disorders, and developmental disorders.

SUMMARY OF THE INVENTION

The invention features substantially purified polypeptides, proteins regulating gene expression, referred to collectively as "PRGE" and individually as "PRGE-1", "PRGE-2", "PRGE-3", "PRGE-4", "PRGE-5", "PRGE-6", "PRGE-7", "PRGE-8", "PRGE-9", "PRGE-10", "PRGE-11", "PRGE-12", "PRGE-13", "PRGE-14", "PRGE-15", "PRGE-16", "PRGE-17", "PRGE-18", "PRGE-19", "PRGE-20", "PRGE-21", "PRGE-22", "PRGE-23", "PRGE-24", "PRGE-25", "PRGE-26", "PRGE-27", "PRGE-28", "PRGE-29", "PRGE-30",and "PRGE-31". In one aspect, the invention provides a substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31 (SEQ ID NO:1-31), and fragments thereof.

The invention further provides a substantially purified variant having at least 90% amino acid identity to at least one of the amino acid sequences selected from the group consisting of SEQ ID NO:1-31, and fragments thereof. The invention also provides an isolated and purified polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-31, and fragments thereof. The invention also includes an isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-31, and fragments thereof.

Additionally, the invention provides an isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-31, and fragments thereof. The invention also provides an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide encoding the polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NO:1-31, and fragments thereof.

The invention also provides an isolated and purified polynucleotide comprising a

polynucleotide sequence selected from the group consisting of SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62 (SEQ ID NO:32-62), and fragments thereof. The invention further provides an isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide sequence selected from the group consisting of SEQ ID NO:32-62, and fragments thereof. The invention also provides an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:32-62, and fragments thereof.

The invention also provides a method for detecting a polynucleotide in a sample containing nucleic acids, the method comprising the steps of (a) hybridizing the complement of the polynucleotide sequence to at least one of the polynucleotides of the sample, thereby forming a hybridization complex; and (b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of a polynucleotide in the sample. In one aspect, the method further comprises amplifying the polynucleotide prior to hybridization.

The invention further provides an expression vector containing at least a fragment of the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-31, and fragments thereof. In another aspect, the expression vector is contained within a host cell.

The invention also provides a method for producing a polypeptide, the method comprising the steps of: (a) culturing the host cell containing an expression vector containing at least a fragment of a polynucleotide under conditions suitable for the expression of the polypeptide; and (b) recovering the polypeptide from the host cell culture.

The invention also provides a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-31, and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention further includes a purified antibody which binds to a polypeptide selected from the group consisting of SEQ ID NO:1-31, and fragments thereof. The invention also provides a purified agonist and a purified antagonist to the polypeptide.

The invention also provides a method for treating or preventing a disorder associated with decreased expression or activity of PRGE, the method comprising administering to a subject in need of such treatment an effective amount of a pharmaceutical composition comprising a

substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-31, and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention also provides a method for treating or preventing a disorder associated with increased expression or activity of PRGE, the method comprising administering to a subject in need of such treatment an effective amount of an antagonist of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-31, and fragments thereof.

BRIEF DESCRIPTION OF THE TABLES

Table 1 shows nucleotide and polypeptide sequence identification numbers (SEQ ID NO), clone identification numbers (clone ID), cDNA libraries, and cDNA fragments used to assemble full-length sequences encoding PRGE.

Table 2 shows features of each polypeptide sequence including potential motifs, homologous sequences, and methods and algorithms used for identification of PRGE.

Table 3 shows the tissue-specific expression patterns of each nucleic acid sequence as determined by northern analysis, diseases, disorders, or conditions associated with these tissues, and the vector into which each cDNA was cloned.

Table 4 describes the tissues used to construct the cDNA libraries from which Incyte cDNA clones encoding PRGE were isolated.

Table 5 shows the programs, their descriptions, references, and threshold parameters used to analyze PRGE.

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention

belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

"PRGE" refers to the amino acid sequences of substantially purified PRGE obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and preferably the human species, from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which, when bound to PRGE, increases or prolongs the duration of the effect of PRGE. Agonists may include proteins, nucleic acids, carbohydrates, or any other molecules which bind to and modulate the effect of PRGE.

An "allelic variant" is an alternative form of the gene encoding PRGE. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. Any given natural or recombinant gene may have none, one, or many allelic forms. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding PRGE include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polynucleotide the same as PRGE or a polypeptide with at least one functional characteristic of PRGE. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding PRGE, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding PRGE. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent PRGE. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of PRGE is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, positively charged amino acids may include lysine and arginine, and amino acids with

uncharged polar head groups having similar hydrophilicity values may include leucine, isoleucine, and valine; glycine and alanine; asparagine and glutamine; serine and threonine; and phenylalanine and tyrosine.

The terms "amino acid" or "amino acid sequence" refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. In this context, "fragments," "immunogenic fragments," or "antigenic fragments" refer to fragments of PRGE which are preferably at least 5 to about 15 amino acids in length, most preferably at least 14 amino acids, and which retain some biological activity or immunological activity of PRGE. Where "amino acid sequence" is recited to refer to an amino acid sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term "antagonist" refers to a molecule which, when bound to PRGE, decreases the amount or the duration of the effect of the biological or immunological activity of PRGE. Antagonists may include proteins, nucleic acids, carbohydrates, antibodies, or any other molecules which decrease the effect of PRGE.

The term "antibody" refers to intact molecules as well as to fragments thereof, such as Fab, $F(ab')_2$, and Fv fragments, which are capable of binding the epitopic determinant. Antibodies that bind PRGE polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that fragment of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (given regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition containing a nucleic acid sequence which is complementary to the "sense" strand of a specific nucleic acid sequence. Antisense molecules

may be produced by any method including synthesis or transcription. Once introduced into a cell, the complementary nucleotides combine with natural sequences produced by the cell to form duplexes and to block either transcription or translation. The designation "negative" can refer to the antisense strand, and the designation "positive" can refer to the sense strand.

The term "biologically active," refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" refers to the capability of the natural, recombinant, or synthetic PRGE, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The terms "complementary" or "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence "5' A-G-T 3'" bonds to the complementary sequence "3' T-C-A 5'." Complementarity between two single-stranded molecules may be "partial," such that only some of the nucleic acids bind, or it may be "complete," such that total complementarity exists between the single stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands. This is of particular importance in amplification reactions, which depend upon binding between nucleic acids strands, and in the design and use of peptide nucleic acid (PNA) molecules.

A "composition comprising a given polynucleotide sequence" or a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding PRGE or fragments of PRGE may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence"refers to a nucleic acid sequence which has been resequenced to resolve uncalled bases, extended using XL-PCR kit (Perkin-Elmer, Norwalk CT) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from the overlapping sequences of more than one Incyte Clone using a computer program for fragment assembly, such as the GELVIEW Fragment Assembly system (GCG, Madison WI). Some sequences have been both extended and assembled to produce the consensus sequence.

The term "correlates with expression of a polynucleotide" indicates that the detection of the presence of nucleic acids, the same or related to a nucleic acid sequence encoding PRGE, by

northern analysis is indicative of the presence of nucleic acids encoding PRGE in a sample, and thereby correlates with expression of the transcript from the polynucleotide encoding PRGE.

A "deletion"refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to the chemical modification of a polypeptide sequence, or a polynucleotide sequence. Chemical modifications of a polynucleotide sequence can include, for example, replacement of hydrogen by an alkyl, acyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

The term "similarity" refers to a degree of complementarity. There may be partial similarity or complete similarity. The word "identity" may substitute for the word "similarity." A partially complementary sequence that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid is referred to as "substantially similar." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or northern blot, solution hybridization, and the like) under conditions of reduced stringency. A substantially similar sequence or hybridization probe will compete for and inhibit the binding of a completely similar (identical) sequence to the target sequence under conditions of reduced stringency. This is not to say that conditions of reduced stringency are such that non-specific binding is permitted, as reduced stringency conditions require that the binding of two sequences to one another be a specific (i.e., a selective) interaction. The absence of non-specific binding may be tested by the use of a second target sequence which lacks even a partial degree of complementarity (e.g., less than about 30% similarity or identity). In the absence of non-specific binding, the substantially similar sequence or probe will not hybridize to the second non-complementary target sequence.

The phrases "percent identity" or "% identity" refer to the percentage of sequence similarity found in a comparison of two or more amino acid or nucleic acid sequences. Percent identity can be determined electronically, e.g., by using the MEGALIGN program (DNASTAR, Madison WI) which creates alignments between two or more sequences according to methods selected by the user, e.g., the clustal method. (See, e.g., Higgins, D.G. and P.M. Sharp (1988) Gene 73:237-244.) The clustal algorithm groups sequences into clusters by examining the distances between all pairs. The clusters are aligned pairwise and then in groups. The percentage similarity between two amino acid sequences, e.g., sequence A and sequence B, is calculated by dividing the length of sequence A, minus the number of gap residues in sequence A, minus the

number of gap residues in sequence B, into the sum of the residue matches between sequence A and sequence B, times one hundred. Gaps of low or of no similarity between the two amino acid sequences are not included in determining percentage similarity. Percent identity between nucleic acid sequences can also be counted or calculated by other methods known in the art, e.g., the Jotun Hein method. (See, e.g., Hein, J. (1990) Methods Enzymol. 183:626-645.) Identity between sequences can also be determined by other methods known in the art, e.g., by varying hybridization conditions.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements required for stable mitotic chromosome segregation and maintenance.

The term "humanized antibody" refers to antibody molecules in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to any process by which a strand of nucleic acid binds with a complementary strand through base pairing.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" or "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively, to the sequence found in the naturally occurring molecule.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

The term "microarray" refers to an arrangement of distinct polynucleotides on a substrate.

The terms "element" or "array element" in a microarray context, refer to hybridizable polynucleotides arranged on the surface of a substrate.

The term "modulate" refers to a change in the activity of PRGE. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of PRGE.

The phrases "nucleic acid" or "nucleic acid sequence," as used herein, refer to a

nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material. In this context, "fragments" refers to those nucleic acid sequences which, comprise a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:32-62, for example, as distinct from any other sequence in the same genome. For example, a fragment of SEQ ID NO:32-62 is useful in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:32-62 from related polynucleotide sequences. A fragment of SEQ ID NO:32-62 is at least about 15-20 nucleotides in length. The precise length of the fragment of SEQ ID NO:32-62 and the region of SEQ ID NO:32-62 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment. Alternatively, a fragment when translated, would produce polypeptides retaining some functional characteristic, e.g., antigenicity, or structural domain characteristic, e.g., ATP-binding site, of the full-length polypeptide.

The terms "operably associated" or "operably linked" refer to functionally related nucleic acid sequences. A promoter is operably associated or operably linked with a coding sequence if the promoter controls the translation of the encoded polypeptide. While operably associated or operably linked nucleic acid sequences can be contiguous and in the same reading frame, certain genetic elements, e.g., repressor genes, are not contiguously linked to the sequence encoding the polypeptide but still bind to operator sequences that control expression of the polypeptide.

The term "oligonucleotide" refers to a nucleic acid sequence of at least about 6 nucleotides to 60 nucleotides, preferably about 15 to 30 nucleotides, and most preferably about 20 to 25 nucleotides, which can be used in PCR amplification or in a hybridization assay or microarray. "Oligonucleotide" is substantially equivalent to the terms "amplimer," "primer," "oligomer," and "probe," as these terms are commonly defined in the art.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

The term "sample" is used in its broadest sense. A sample suspected of containing nucleic acids encoding PRGE, or fragments thereof, or PRGE itself, may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" or "specifically binding" refer to that interaction between a

protein or peptide and an agonist, an antibody, or an antagonist. The interaction is dependent upon the presence

of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide containing the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "stringent conditions" refers to conditions which permit hybridization between polynucleotides and the claimed polynucleotides. Stringent conditions can be defined by salt concentration, the concentration of organic solvent, e.g., formamide, temperature, and other conditions well known in the art. In particular, stringency can be increased by reducing the concentration of salt, increasing the concentration of formamide, or raising the hybridization temperature.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least about 60% free, preferably about 75% free, and most preferably about 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acids or nucleotides by different amino acids or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

"Transformation" describes a process by which exogenous DNA enters and changes a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed" cells includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "variant" of PRGE polypeptides refers to an amino acid sequence that is altered by one

or more amino acid residues. The variant may have "conservative" changes, wherein a substituted amino acid has similar structural or chemical properties (e.g., replacement of leucine with isoleucine). More rarely, a variant may have "nonconservative" changes (e.g., replacement of glycine with tryptophan). Analogous minor variations may also include amino acid deletions or insertions, or both. Guidance in determining which amino acid residues may be substituted, inserted, or deleted without abolishing biological or immunological activity may be found using computer programs well known in the art, for example, LASERGENE software (DNASTAR).

The term "variant," when used in the context of a polynucleotide sequence, may encompass a polynucleotide sequence related to PRGE. This definition may also include, for example, "allelic" (as defined above), "splice," "species," or "polymorphic" variants. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or an absence of domains. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

THE INVENTION

The invention is based on the discovery of new human proteins regulating gene expression (PRGE), the polynucleotides encoding PRGE, and the use of these compositions for the diagnosis, treatment, or prevention of reproductive disorders, nervous disorders, and diseases associated with cell proliferation and differentiation, including cancer, immune disorders, and developmental disorders.

Table 1 lists the Incyte Clones used to derive full length nucleotide sequences encoding PRGE. Columns 1 and 2 show the sequence identification numbers (SEQ ID NO) of the amino acid and nucleic acid sequences, respectively. Column 3 shows the Clone ID of the Incyte Clone in which nucleic acids encoding each PRGE were identified, and column 4, the cDNA libraries from which these clones were isolated. Column 5 shows Incyte clones, their corresponding cDNA libraries, and shotgun sequences. The clones and shotgun sequences are part of the consensus nucleotide sequence of each PRGE. The regions of the full-length nucleotide sequence of each PRGE to which the clones and shotgun sequences correspond are listed in Column 5. The clones and fragments are useful as fragments in hybridization technologies.

The columns of Table 2 show various properties of the polypeptides of the invention: column 1 references the SEQ ID NO; column 2 shows the number of amino acid residues in each polypeptide; column 3, potential phosphorylation sites; column 4, potential glycosylation sites; column 5, the amino acid residues comprising signature sequences and motifs; column 6, the identity of each protein; and column 7, analytical methods used to identify each protein through sequence homology and protein motifs.

The columns of Table 3 show the tissue-specificity and diseases, disorders, or conditions associated with nucleotide sequences encoding PRGE. The first column of Table 3 lists the nucleotide sequence identifiers. The second column lists tissue categories which express PRGE as a fraction of total tissue categories expressing PRGE. The third column lists the diseases, disorders, or conditions associated with those tissues expressing PRGE. The fourth column lists the vectors used to subclone the cDNA library.

The following fragments of the nucleotide sequences encoding PRGE are useful in hybridization or amplification technologies to identify SEQ ID NO:56-62 and to distinguish between SEQ ID NO:56-62 and related polynucleotide sequences. The useful fragments are the fragment of SEQ ID NO:56 from about nucleotide 1675 to about nucleotide 1719; the fragment of SEQ ID NO:57 from about nucleotide 379 to about nucleotide 423; the fragment of SEQ ID NO:58 from about nucleotide 596 to about nucleotide 640; the fragment of SEQ ID NO:59 from about nucleotide 219 to about nucleotide 263; the fragment of SEQ ID NO:60 from about nucleotide 732 to about nucleotide 776; the fragment of SEQ ID NO:61 from about nucleotide 197 to about nucleotide 244; and the fragment of SEQ ID NO:62 from about nucleotide 217 to about nucleotide 261.

The invention also encompasses PRGE variants. A preferred PRGE variant is one which has at least about 80%, more preferably at least about 90%, and most preferably at least about 95% amino acid sequence identity to the PRGE amino acid sequence, and which contains at least one functional or structural characteristic of PRGE.

The invention also encompasses polynucleotides which encode PRGE. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:32-62, which encodes PRGE.

The invention also encompasses a variant of a polynucleotide sequence encoding PRGE. In particular, such a variant polynucleotide sequence will have at least about 80%, more preferably at least about 90%, and most preferably at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding PRGE. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:32-62 which has at least about 80%, more preferably at least about 90%, and most

preferably at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:32-62. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of PRGE.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding PRGE, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring PRGE, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode PRGE and its variants are preferably capable of hybridizing to the nucleotide sequence of the naturally occurring PRGE under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding PRGE or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding PRGE and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode PRGE and PRGE derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding PRGE or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:32-62 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) Methods Enzymol. 152:399-407; Kimmel, A.R. (1987) Methods Enzymol. 152:507-511.) For example, stringent salt concentration will ordinarily be less than about 750 mM NaCl and 75 mM trisodium citrate, preferably less than about 500 mM NaCl and 50 mM trisodium citrate, and most preferably less than about 250 mM NaCl and 25 mM trisodium citrate. Low

stringency hybridization can be obtained in the absence of organic solvent, e.g., formamide, while high stringency hybridization can be obtained in the presence of at least about 35% formamide, and most preferably at least about 50% formamide. Stringent temperature conditions will ordinarily include temperatures of at least about 30°C, more preferably of at least about 37°C, and most preferably of at least about 42°C. Varying additional parameters, such as hybridization time, the concentration of detergent, e.g., sodium dodecyl sulfate (SDS), and the inclusion or exclusion of carrier DNA, are well known to those skilled in the art. Various levels of stringency are accomplished by combining these various conditions as needed. In a preferred embodiment, hybridization will occur at 30°C in 750 mM NaCl, 75 mM trisodium citrate, and 1% SDS. In a more preferred embodiment, hybridization will occur at 37°C in 500 mM NaCl, 50 mM trisodium citrate, 1% SDS, 35% formamide, and 100 μg/ml denatured salmon sperm DNA (ssDNA). In a most preferred embodiment, hybridization will occur at 42°C in 250 mM NaCl, 25 mM trisodium citrate, 1% SDS, 50 % formamide, and 200 μg/ml ssDNA. Useful variations on these conditions will be readily apparent to those skilled in the art.

The washing steps which follow hybridization can also vary in stringency. Wash stringency conditions can be defined by salt concentration and by temperature. As above, wash stringency can be increased by decreasing salt concentration or by increasing temperature. For example, stringent salt concentration for the wash steps will preferably be less than about 30 mM NaCl and 3 mM trisodium citrate, and most preferably less than about 15 mM NaCl and 1.5 mM trisodium citrate. Stringent temperature conditions for the wash steps will ordinarily include temperature of at least about 25°C, more preferably of at least about 42°C, and most preferably of at least about 68°C. In a preferred embodiment, wash steps will occur at 25°C in 30 mM NaCl, 3 mM trisodium citrate, and 0.1% SDS. In a more preferred embodiment, wash steps will occur at 42°C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. In a most preferred embodiment, wash steps will occur at 68°C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. Additional variations on these conditions will be readily apparent to those skilled in the art.

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Perkin-Elmer), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the Hamilton MICROLAB 2200 (Hamilton, Reno NV), Peltier Thermal Cycler 200 (PTC200; MJ Research, Watertown MA) and the ABI CATALYST 800 (Perkin-Elmer). Sequencing is then carried out using either ABI 373 or 377

DNA sequencing systems (Perkin-Elmer) or the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA). The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding PRGE may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Applic. 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-306). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for

detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Perkin-Elmer), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode PRGE may be cloned in recombinant DNA molecules that direct expression of PRGE, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express PRGE.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter PRGE-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

In another embodiment, sequences encoding PRGE may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucl. Acids Res. Symp. Ser. 215-223, and Horn, T. et al. (1980) Nucl. Acids Res. Symp. Ser. 225-232.) Alternatively, PRGE itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solid-phase techniques. (See, e.g., Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the ABI 431A Peptide Synthesizer (Perkin-Elmer). Additionally, the amino acid sequence of PRGE, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g, Chiez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY.)

In order to express a biologically active PRGE, the nucleotide sequences encoding PRGE or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which

contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding PRGE. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding PRGE. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding PRGE and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding PRGE and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding PRGE. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus,TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding PRGE. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding PRGE can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or pSPORT1 plasmid (Life Technologies). Ligation of sequences encoding PRGE into the

vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of PRGE are needed, e.g. for the production of antibodies, vectors which direct high level expression of PRGE may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of PRGE. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra; Grant et al. (1987) Methods Enzymol. 153:516-54; and Scorer, C. A. et al. (1994) Bio/Technology 12:181-184.)

Plant systems may also be used for expression of PRGE. Transcription of sequences encoding PRGE may be driven viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; and Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding PRGE may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses PRGE in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino

polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat Genet. 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of PRGE in cell lines is preferred. For example, sequences encoding PRGE can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk*$^-$ or *apr*$^-$ cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides, neomycin and G-418; and *als* or *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), ß glucuronidase and its substrate ß-glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding PRGE is inserted within a marker gene sequence, transformed cells containing sequences encoding PRGE can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding PRGE under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding PRGE and that

express PRGE may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of PRGE using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on PRGE is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St Paul MN, Sect. IV; Coligan, J. E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ).

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding PRGE include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding PRGE, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding PRGE may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode PRGE may be designed to contain signal sequences which direct secretion of PRGE through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the

inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38), are available from the American Type Culture Collection (ATCC, Bethesda MD) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding PRGE may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric PRGE protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of PRGE activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the PRGE encoding sequence and the heterologous protein sequence, so that PRGE may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled PRGE may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract systems (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, preferably ^{35}S-methionine.

Fragments of PRGE may be produced not only by recombinant production, but also by direct peptide synthesis using solid-phase techniques. (See, e.g., Creighton, supra, pp. 55-60.) Protein synthesis may be performed by manual techniques or by automation. Automated synthesis may be achieved, for example, using the ABI 431A Peptide Synthesizer (Perkin-Elmer). Various

fragments of PRGE may be synthesized separately and then combined to produce the full length molecule.

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of PRGE and proteins regulating gene expression. In addition, the expression of PRGE is closely associated with cancer and other cell proliferative conditions, differentiated cells, inflammation and the immune response, and is found in reproductive and nervous system tissues. Therefore, PRGE appears to play a role in reproductive disorders, nervous disorders, and diseases associated with cell proliferation and differentiation, including cancer, immune disorders, and developmental disorders. In the treatment of the above conditions associated with increased PRGE expression or activity, it is desirable to decrease the expression or activity of PRGE. In the treatment of the above conditions associated with decreased PRGE expression or activity, it is desirable to increase the expression or activity of PRGE.

Therefore, in one embodiment, PRGE or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of PRGE. Examples of such a disorder include, but are not limited to, a reproductive disorder such as disorders of prolactin production; infertility, including tubal disease, ovulatory defects, and endometriosis; disruptions of the estrous cycle, disruptions of the menstrual cycle, polycystic ovary syndrome, ovarian hyperstimulation syndrome, endometrial and ovarian tumors, uterine fibroids, autoimmune disorders, ectopic pregnancies, and teratogenesis; cancer of the breast, fibrocystic breast disease, and galactorrhea; disruptions of spermatogenesis, abnormal sperm physiology, cancer of the testis, cancer of the prostate, benign prostatic hyperplasia, prostatitis, Peyronie's disease, impotence, carcinoma of the male breast, and gynecomastia; and a nervous disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease; prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome; fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other

neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis; inherited, metabolic, endocrine, and toxic myopathies; myasthenia gravis, periodic paralysis; mental disorders including mood, anxiety, and schizophrenic disorders; akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, and Tourette's disorder; a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia; cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an immune disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, trauma, X-linked agammaglobinemia of Bruton, common variable immunodeficiency (CVI), DiGeorge's syndrome (thymic hypoplasia), thymic dysplasia, isolated IgA deficiency, severe combined immunodeficiency disease (SCID), immunodeficiency with thrombocytopenia and eczema (Wiskott-Aldrich syndrome), Chediak-Higashi syndrome, chronic granulomatous diseases, hereditary angioneurotic edema, and immunodeficiency associated with Cushing's disease; and a developmental disorder such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucoepithelial dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Syndenham's

chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, sensorineural hearing loss, and any disorder associated with cell growth and differentiation, embryogenesis, and morphogenesis involving any tissue, organ, or system of a subject, e.g., the brain, adrenal gland, kidney, skeletal or reproductive system.

In another embodiment, a vector capable of expressing PRGE or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of PRGE including, but not limited to, those described above.

In a further embodiment, a pharmaceutical composition comprising a substantially purified PRGE in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of PRGE including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of PRGE may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of PRGE including, but not limited to, those listed above.

In a further embodiment, an antagonist of PRGE may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of PRGE. Examples of such disorders include, but are not limited to, those described above. In one aspect, an antibody which specifically binds PRGE may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissue which express PRGE.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding PRGE may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of PRGE including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of PRGE may be produced using methods which are generally known in the art. In particular, purified PRGE may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind PRGE. Antibodies to PRGE may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments,

and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are especially preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with PRGE or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to PRGE have an amino acid sequence consisting of at least about 5 amino acids, and, more preferably, of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein and contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of PRGE amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to PRGE may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. 80:2026-2030; and Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce PRGE-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton D.R. (1991) Proc. Natl. Acad. Sci. 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. 86:

3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for PRGE may also be generated. For example, such fragments include, but are not limited to, F(ab')2 fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between PRGE and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering PRGE epitopes is preferred, but a competitive binding assay may also be employed (Pound, supra).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for PRGE. Affinity is expressed as an association constant, K_a, which is defined as the molar concentration of PRGE-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple PRGE epitopes, represents the average affinity, or avidity, of the antibodies for PRGE. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular PRGE epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} L/mole are preferred for use in immunoassays in which the PRGE-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7 L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of PRGE, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington, DC; Liddell, J. E. and Cryer, A. (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is preferred for use in procedures requiring precipitation of PRGE-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available.

(See, e.g., Catty, supra, and Coligan et al. supra.)

In another embodiment of the invention, the polynucleotides encoding PRGE, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, the complement of the polynucleotide encoding PRGE may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding PRGE. Thus, complementary molecules or fragments may be used to modulate PRGE activity, or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding PRGE.

Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. Methods which are well known to those skilled in the art can be used to construct vectors to express nucleic acid sequences complementary to the polynucleotides encoding PRGE. (See, e.g., Sambrook, supra; Ausubel, 1995, supra.)

Genes encoding PRGE can be turned off by transforming a cell or tissue with expression vectors which express high levels of a polynucleotide, or fragment thereof, encoding PRGE. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector, and may last even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing complementary sequences or antisense molecules (DNA, RNA, or PNA) to the control, 5', or regulatory regions of the gene encoding PRGE. Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, are preferred. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the

ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding PRGE.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding PRGE. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nature Biotechnology 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as dogs, cats, cows, horses, rabbits, monkeys, and most preferably, humans.

An additional embodiment of the invention relates to the administration of a pharmaceutical or sterile composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of PRGE, antibodies to PRGE, and mimetics, agonists, antagonists, or inhibitors of PRGE. The compositions may be administered alone or in combination with at least one other agent, such as a stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs, or hormones.

The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA).

Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combining active compounds with solid excipient and processing the resultant mixture of granules (optionally, after grinding) to obtain tablets or dragee cores. Suitable auxiliaries can be added, if desired. Suitable excipients include carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, and sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums, including arabic and tragacanth; and proteins, such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, and alginic acid or a salt thereof, such as sodium alginate.

Dragee cores may be used in conjunction with suitable coatings, such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for

product identification or to characterize the quantity of active compound, i.e., dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with fillers or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils, such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate, triglycerides, or liposomes. Non-lipid polycationic amino polymers may also be used for delivery. Optionally, the suspension may also contain suitable stabilizers or agents to increase the solubility of the compounds and allow for the preparation of highly concentrated solutions.

For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention may be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

The pharmaceutical composition may be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, and succinic acid. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preferred preparation may be a lyophilized powder which may contain any or all of the following: 1 mM to 50 mM histidine, 0.1% to 2% sucrose, and 2% to 7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of PRGE, such labeling would include amount, frequency, and method of administration.

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the

art.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells or in animal models such as mice, rats, rabbits, dogs, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example PRGE or fragments thereof, antibodies of PRGE, and agonists, antagonists or inhibitors of PRGE, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, and it can be expressed as the LD_{50}/ED_{50} ratio. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 μg to 100,000 μg, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind PRGE may be used for the diagnosis of disorders characterized by expression of PRGE, or in assays to monitor patients being

treated with PRGE or agonists, antagonists, or inhibitors of PRGE. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for PRGE include methods which utilize the antibody and a label to detect PRGE in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring PRGE, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of PRGE expression. Normal or standard values for PRGE expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, preferably human, with antibody to PRGE under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, preferably by photometric means. Quantities of PRGE expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding PRGE may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantitate gene expression in biopsied tissues in which expression of PRGE may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of PRGE, and to monitor regulation of PRGE levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding PRGE or closely related molecules may be used to identify nucleic acid sequences which encode PRGE. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification (maximal, high, intermediate, or low), will determine whether the probe identifies only naturally occurring sequences encoding PRGE, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and should preferably have at least 50% sequence identity to any of the PRGE encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:32-62 or from genomic sequences including promoters, enhancers, and introns of the PRGE gene.

Means for producing specific hybridization probes for DNAs encoding PRGE include the cloning of polynucleotide sequences encoding PRGE or PRGE derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding PRGE may be used for the diagnosis of disorders associated with expression of PRGE. Examples of such a disorder include, but are not limited to, a reproductive disorder such as disorders of prolactin production; infertility, including tubal disease, ovulatory defects, and endometriosis; disruptions of the estrous cycle, disruptions of the menstrual cycle, polycystic ovary syndrome, ovarian hyperstimulation syndrome, endometrial and ovarian tumors, uterine fibroids, autoimmune disorders, ectopic pregnancies, and teratogenesis; cancer of the breast, fibrocystic breast disease, and galactorrhea; disruptions of spermatogenesis, abnormal sperm physiology, cancer of the testis, cancer of the prostate, benign prostatic hyperplasia, prostatitis, Peyronie's disease, impotence, carcinoma of the male breast, and gynecomastia; and a nervous disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease; prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome; fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis; inherited, metabolic, endocrine, and toxic myopathies; myasthenia gravis, periodic paralysis; mental disorders including mood, anxiety, and schizophrenic disorders; akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, and Tourette's disorder; a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis,

primary thrombocythemia; cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an immune disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, trauma, X-linked agammaglobinemia of Bruton, common variable immunodeficiency (CVI), DiGeorge's syndrome (thymic hypoplasia), thymic dysplasia, isolated IgA deficiency, severe combined immunodeficiency disease (SCID), immunodeficiency with thrombocytopenia and eczema (Wiskott-Aldrich syndrome), Chediak-Higashi syndrome, chronic granulomatous diseases, hereditary angioneurotic edema, and immunodeficiency associated with Cushing's disease; and a developmental disorder such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucoepithelial dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Syndenham's chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, sensorineural hearing loss, and any disorder associated with cell growth and differentiation, embryogenesis, and morphogenesis involving any tissue, organ, or system of a subject, e.g., the brain, adrenal gland, kidney, skeletal or reproductive system.

The polynucleotide sequences encoding PRGE may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to

detect altered PRGE expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding PRGE may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding PRGE may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantitated and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding PRGE in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of PRGE, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding PRGE, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding PRGE may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a

polynucleotide encoding PRGE, or a fragment of a polynucleotide complementary to the polynucleotide encoding PRGE, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantitation of closely related DNA or RNA sequences.

Methods which may also be used to quantitate the expression of PRGE include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in an ELISA format where the oligomer of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as targets in a microarray. The microarray can be used to monitor the expression level of large numbers of genes simultaneously and to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, and to develop and monitor the activities of therapeutic agents.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.)

In another embodiment of the invention, nucleic acid sequences encoding PRGE may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical chromosome mapping techniques and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) site. Correlation between the location of the gene encoding PRGE on a physical chromosomal map and a specific disorder, or a

predisposition to a specific disorder, may help define the region of DNA associated with that disorder. The nucleotide sequences of the invention may be used to detect differences in gene sequences among normal, carrier, and affected individuals.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, PRGE, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between PRGE and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with PRGE, or fragments thereof, and washed. Bound PRGE is then detected by methods well known in the art. Purified PRGE can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding PRGE specifically compete with a test compound for binding PRGE. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with PRGE.

In additional embodiments, the nucleotide sequences which encode PRGE may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such

properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, in particular U.S. Ser. No. 60/089,029, U.S. Ser. No. 60/094,575, and U.S. Ser. No. 60/104,624, are hereby expressly incorporated by reference.

EXAMPLES

I. Construction of cDNA Libraries

RNA was purchased from Clontech or isolated from tissues described in Table 4. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A+) RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Valencia CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPT plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra, units 5.1-6.6). Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), pSPORT1 plasmid (Life Technologies), or pINCY (Incyte Pharmaceuticals, Palo Alto CA). Recombinant plasmids

were transformed into competent E. coli cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5α, DH10B, or ElectroMAX DH10B from Life Technologies. **II.**

Isolation of cDNA Clones

Plasmids were recovered from host cells by in vivo excision, using the UNIZAP vector system (Stratagene) or cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the REAL Prep 96 plasmid kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a Fluoroskan II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

The cDNAs were prepared for sequencing using the ABI CATALYST 800 (Perkin-Elmer) or the HYDRA microdispenser (Robbins Scientific) or MICROLAB 2200 (Hamilton) systems in combination with the PTC-200 thermal cyclers (MJ Research). The cDNAs were sequenced using the ABI PRISM 373 or 377 sequencing systems (Perkin-Elmer) and standard ABI protocols, base calling software, and kits. In one alternative, cDNAs were sequenced using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics). In another alternative, the cDNAs were amplified and sequenced using the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer). In yet another alternative, cDNAs were sequenced using solutions and dyes from Amersham Pharmacia Biotech. Reading frames for the ESTs were determined using standard methods (reviewed in Ausubel, 1997, supra, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example V.

The polynucleotide sequences derived from cDNA, extension, and shotgun sequencing were assembled and analyzed using a combination of software programs which utilize algorithms well known to those skilled in the art. Table 5 summarizes the software programs, descriptions, references, and threshold parameters used. The first column of Table 5 shows the tools, programs, and algorithms used, the second column provides a brief description thereof, the third column presents the references which are incorporated by reference herein, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate

the strength of a match between two sequences (the higher the probability the greater the homology). Sequences were analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR).

The polynucleotide sequences were validated by removing vector, linker, and polyA sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programing, and dinucleotide nearest neighbor analysis. The sequences were then queried against a selection of public databases such as GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS to acquire annotation, using programs based on BLAST, FASTA, and BLIMPS. The sequences were assembled into full length polynucleotide sequences using programs based on Phred, Phrap, and Consed, and were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length amino acid sequences, and these full length sequences were subsequently analyzed by querying against databases such as the GenBank databases (described above), SwissProt, BLOCKS, PRINTS, Prosite, and Hidden Markov Model (HMM)-based protein family databases such as PFAM. HMM is a probalistic approach which analyzes consensus primary structures of gene families. (See, e.g., Eddy, S.R. (1996) Cur. Opin. Str. Biol. 6:361-365.)

The programs described above for the assembly and analysis of full length polynucleotide and amino acid sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:32-62. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies were described in The Invention section above.

IV. Northern Analysis

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, supra, ch. 7; Ausubel, 1995, supra, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in nucleotide databases such as GenBank or LIFESEQ database (Incyte Pharmaceuticals). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\%\ \text{sequence identity} \times \%\ \text{maximum BLAST score}}{100}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact

within a 1% to 2% error, and, with a product score of 70, the match will be exact. Similar molecules are usually identified by selecting those which show product scores between 15 and 40, although lower scores may identify related molecules.

The results of northern analyses are reported as a percentage distribution of libraries in which the transcript encoding PRGE occurred. Analysis involved the categorization of cDNA libraries by organ/tissue and disease. The organ/tissue categories included cardiovascular, dermatologic, developmental, endocrine, gastrointestinal, hematopoietic/immune, musculoskeletal, nervous, reproductive, and urologic. The disease/condition categories included cancer, inflammation/trauma, cell proliferation, neurological, and pooled. For each category, the number of libraries expressing the sequence of interest was counted and divided by the total number of libraries across all categories. Percentage values of tissue-specific and disease- or condition-specific expression are reported in Table 3.

V. Extension of PRGE Encoding Polynucleotides

The full length nucleic acid sequences of SEQ ID NO:56-62 were produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer, to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+}, $(NH_4)_2SO_4$, and β-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 µl PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 µl of undiluted PCR product into each well of an opaque fluorimeter

plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 μl to 10 μl aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose mini-gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent E. coli cells. Transformed cells were selected on antibiotic-containing media, individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethysulphoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer).

In like manner, the nucleotide sequences of SEQ ID NO:56-62 are used to obtain 5' regulatory sequences using the procedure above, oligonucleotides designed for such extension, and an appropriate genomic library.

The nucleic acid sequences of SEQ ID NO:32-55 were used to design oligonucleotide primers for extending a partial nucleotide sequence to full length. For each nucleic acid sequence, one primer was synthesized to initiate extension of an antisense polynucleotide, and the other was synthesized to initiate extension of a sense polynucleotide. Primers were used to facilitate the extension of the known sequence "outward" generating amplicons containing new unknown nucleotide sequence for the region of interest. The initial primers were designed from the cDNA using OLIGO™ 4.06 (National Biosciences, Plymouth, MN), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides

which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries (GIBCO BRL) were used to extend the sequence. If more than one extension is necessary or desired, additional sets of primers are designed to further extend the known region.

High fidelity amplification was obtained by following the instructions for the XL-PCR™ kit (Perkin Elmer) and thoroughly mixing the enzyme and reaction mix. PCR was performed using the Peltier Thermal Cycler (PTC200; M.J. Research, Watertown, MA), beginning with 40 pmol of each primer and the recommended concentrations of all other components of the kit, with the following parameters:

Step 1	94° C for 1 min (initial denaturation)
Step 2	65° C for 1 min
Step 3	68° C for 6 min
Step 4	94° C for 15 sec
Step 5	65° C for 1 min
Step 6	68° C for 7 min
Step 7	Repeat steps 4 through 6 for an additional 15 cycles
Step 8	94° C for 15 sec
Step 9	65° C for 1 min
Step 10	68° C for 7:15 min
Step 11	Repeat steps 8 through 10 for an additional 12 cycles
Step 12	72° C for 8 min
Step 13	4° C (and holding)

A 5 μl to 10 μl aliquot of the reaction mixture was analyzed by electrophoresis on a low concentration (about 0.6% to 0.8%) agarose mini-gel to determine which reactions were successful in extending the sequence. Bands thought to contain the largest products were excised from the gel, purified using QIAQUICK™ (QIAGEN Inc.), and trimmed of overhangs using Klenow enzyme to facilitate religation and cloning.

After ethanol precipitation, the products were redissolved in 13 μl of ligation buffer, 1μl T4-DNA ligase (15 units) and 1μl T4 polynucleotide kinase were added, and the mixture was incubated at room temperature for 2 to 3 hours, or overnight at 16° C. Competent E. coli cells (in 40 μl of appropriate media) were transformed with 3 μl of ligation mixture and cultured in 80 μl of SOC medium. (See, e.g., Sambrook, supra, Appendix A, p. 2.) After incubation for one hour at 37°C, the E. coli mixture was plated on Luria Bertani (LB) agar (See, e.g., Sambrook, supra, Appendix A, p. 1) containing carbenicillin (2x carb). The following day, several colonies were randomly picked from each plate and cultured in 150 μl of liquid LB/2x carb medium placed in an individual well of an appropriate commercially-available sterile 96-well microtiter plate. The following day, 5 μl of each overnight culture was transferred into a non-sterile 96-well plate and, after dilution 1:10 with water, 5 μl from each sample was transferred into a PCR array.

For PCR amplification, 18 μl of concentrated PCR reaction mix (3.3x) containing 4 units

of rTth DNA polymerase, a vector primer, and one or both of the gene specific primers used for the extension reaction were added to each well. Amplification was performed using the following conditions:

Step 1	94° C for 60 sec
Step 2	94° C for 20 sec
Step 3	55° C for 30 sec
Step 4	72° C for 90 sec
Step 5	Repeat steps 2 through 4 for an additional 29 cycles
Step 6	72° C for 180 sec
Step 7	4° C (and holding)

Aliquots of the PCR reactions were run on agarose gels together with molecular weight markers. The sizes of the PCR products were compared to the original partial cDNAs, and appropriate clones were selected, ligated into plasmid, and sequenced.

In like manner, the nucleotide sequences of SEQ ID NO:32-55 are used to obtain 5' regulatory sequences using the procedure above, oligonucleotides designed for 5' extension, and an appropriate genomic library.

VI. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:32-62 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μCi of [γ-^{32}P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba1, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under increasingly stringent conditions up to 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. After XOMAT-AR film (Eastman Kodak, Rochester NY) is exposed to the blots to film for several hours, hybridization patterns are compared visually.

VII. Microarrays

A chemical coupling procedure and an ink jet device can be used to synthesize array

elements on the surface of a substrate. (See, e.g., Baldeschweiler, supra.) An array analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced by hand or using available methods and machines and contain any appropriate number of elements. After hybridization, nonhybridized probes are removed and a scanner used to determine the levels and patterns of fluorescence. The degree of complementarity and the relative abundance of each probe which hybridizes to an element on the microarray may be assessed through analysis of the scanned images.

Full-length cDNAs, Expressed Sequence Tags (ESTs), or fragments thereof may comprise the elements of the microarray. Fragments suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). Full-length cDNAs, ESTs, or fragments thereof corresponding to one of the nucleotide sequences of the present invention, or selected at random from a cDNA library relevant to the present invention, are arranged on an appropriate substrate, e.g., a glass slide. The cDNA is fixed to the slide using, e.g., UV cross-linking followed by thermal and chemical treatments and subsequent drying. (See, e.g., Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645.) Fluorescent probes are prepared and used for hybridization to the elements on the substrate. The substrate is analyzed by procedures described above.

VIII. Complementary Polynucleotides

Sequences complementary to the PRGE-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring PRGE. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of PRGE. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the PRGE-encoding transcript.

IX. Expression of PRGE

Expression and purification of PRGE is achieved using bacterial or virus-based expression systems. For expression of PRGE in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express PRGE upon induction with isopropyl beta-

D-thiogalactopyranoside (IPTG). Expression of PRGE in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding PRGE by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E. K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, PRGE is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from PRGE at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch 10 and 16). Purified PRGE obtained by these methods can be used directly in the following activity assay.

X. Demonstration of PRGE Activity

PRGE activity is measured by its ability to stimulate transcription of a reporter gene (Liu, H.Y. et al. (1997) EMBO J. 16(17):5289-5298.) The assay entails the use of a well characterized reporter gene construct, $LexA_{op}$-LacZ, that consists of LexA DNA transcriptional control elements ($LexA_{op}$) fused to sequences encoding the E. coli LacZ enzyme. The methods for constructing and expressing fusion genes, introducing them into cells, and measuring LacZ enzyme activity, are well known to those skilled in the art. Sequences encoding PRGE are cloned into a plasmid that directs the synthesis of a fusion protein, LexA-PRGE, consisting of PRGE and a DNA binding domain derived from the LexA transcription factor. The resulting plasmid, encoding a LexA-PRGE fusion protein, is introduced into yeast cells along with a plasmid containing the $LexA_{op}$-LacZ reporter gene. The amount of LacZ enzyme activity associated with LexA-PRGE transfected cells, relative to control cells, is proportional to the amount of transcription stimulated by the PRGE.

XI. Functional Assays

PRGE function is assessed by expressing the sequences encoding PRGE at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT (Life Technologies) and pCR3.1 (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10 μg of recombinant vector are transiently transfected into a human cell line, preferably of endothelial or hematopoietic origin, using either liposome formulations or electroporation. 1-2 μg of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP, and to evaluate properties, for example, their apoptotic state. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M. G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of PRGE on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding PRGE and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding PRGE and other genes of interest can be analyzed by northern analysis or microarray techniques.

XII. Production of PRGE Specific Antibodies

PRGE substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) Methods Enzymol. 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the PRGE amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is

synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

Typically, oligopeptides 15 residues in length are synthesized using an ABI 431A Peptide Synthesizer (Perkin-Elmer) using fmoc-chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide activity by, for example, binding the peptide to plastic, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XIII. Purification of Naturally Occurring PRGE Using Specific Antibodies

Naturally occurring or recombinant PRGE is substantially purified by immunoaffinity chromatography using antibodies specific for PRGE. An immunoaffinity column is constructed by covalently coupling anti-PRGE antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing PRGE are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of PRGE (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/PRGE binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and PRGE is collected.

XIV. Identification of Molecules Which Interact with PRGE

PRGE, or biologically active fragments thereof, are labeled with ^{125}I Bolton-Hunter reagent. (See, e.g., Bolton et al. (1973) Biochem. J. 133:529.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled PRGE, washed, and any wells with labeled PRGE complex are assayed. Data obtained using different concentrations of PRGE are used to calculate values for the number, affinity, and association of PRGE with the candidate molecules.

Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

Table 1

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
1	32	591290	UTRSNOT01	591290H1 (UTRSNOT01) nucleotide 94-324 of SEQ ID NO:32, 1237346R6 (LUNGTUT02) nucleotide 1134-1550 of SEQ ID NO:32, 1306871X12 (PLACNOT02) nucleotide 174-772 of SEQ ID NO:32, 1306871X36R1 (PLACNOT02) nucleotide 501-817 of SEQ ID NO:32, 2269281H1 (UTRSNOT02) nucleotide 1-260 of SEQ ID NO:32
2	33	815856	OVARTUT01	032530R6 (THP1NOB01) nucleotide 659-1091 of SEQ ID NO:33, 815856H1 (OVARTUT01) nucleotide 925-1160 of SEQ ID NO:33, 2845377F6 (DRGLNOT01) nucleotide 1-621 of SEQ ID NO:33, SAUA00500F1 nucleotide 1384-1180 of SEQ ID NO:33, SAUA00928F1 nucleotide 2087-1887 of SEQ ID NO:33, SAUA01832F1 nucleotide 1453-1909 of SEQ ID NO:33, SAUA03625F1 nucleotide 2104-2305 of SEQ ID NO:33
3	34	996352	KIDNTUT01	895625H1 (BRSTNOT05) nucleotide 574-870 of SEQ ID NO:34, 996352H1 (KIDNTUT01) nucleotide 39-260 of SEQ ID NO:34, 3152180R6 (ADRENON04) nucleotide 1-443 of SEQ ID NO:34, 3152180T6 (ADRENON04) nucleotide 836-239 of SEQ ID NO:34
4	35	1273778	TESTTUT02	118141F1 (MUSCNOT01) nucleotide 147-691 of SEQ ID NO:35, 1273778H1 (TESTTUT02) nucleotide 484-610 of SEQ ID NO:35, 3027384F6 (HEARFET02) nucleotide 715-1365 of SEQ ID NO:35, 3400168H1 (UTRSNOT16) nucleotide 1-235 of SEQ ID NO:35
5	36	1509715	LUNGNOT14	1238984H1 (LUNGTUT02) nucleotide 3-249 of SEQ ID NO:36, 1257794T1 (MENITUT03) nucleotide 2395-1832 of SEQ ID NO:36, 1509715CT1 (LUNGNOT14), 1509715H1 (LUNGNOT14) nucleotide 479-654 of SEQ ID NO:36, 1546123R1 (PROSTUT04) nucleotide 2061-2392 of SEQ ID NO:36

Table 1 (cont.)

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
6	37	1676367	BLADNOT05	1637992F6 (UTRSNOT06) nucleotide 1-460 of SEQ ID NO:37, 1676367F6 (BLADNOT05) nucleotide 147-866 of SEQ ID NO:37, 1676367H1 (BLADNOT05) nucleotide 147-371 of SEQ ID NO:37
7	38	1734119	COLNNOT22	1266849H1 (BRAINOT09) nucleotide 1421-1627 of SEQ ID NO:38, 1455720F1 (COLNFET02) nucleotide 818-1276 of SEQ ID NO:38, 1734119F6 (COLNNOT22) nucleotide 257-731 of SEQ ID NO:38, 1734119H1 (COLNNOT22) nucleotide 257-485 of SEQ ID NO:38, 1794147R6 (PROSTUT05) nucleotide 1088-1632 of SEQ ID NO:38, 2012943H1 (TESTNOT03) nucleotide 496-733 of SEQ ID NO:38, 2362634R6 (LUNGFET05) nucleotide 631-1167 of SEQ ID NO:38, 2529952H1 (GBLANOT02) nucleotide 1-243 of SEQ ID NO:38
8	39	1944813	PITUNOT01	1526383F6 (UCMCL5T01) nucleotide 48-651 of SEQ ID NO:39, 1944813H1 (PITUNOT01) nucleotide 1-242 of SEQ ID NO:39, 2343290F6 (TESTTUT02) nucleotide 307-643 of SEQ ID NO:39, 3010932H1 (MUSCNOT07) nucleotide 727-1024 of SEQ ID NO:39
9	40	2683322	SINIUCT01	1343026F6 (COLNTUT03) nucleotide 1583-1787 of SEQ ID NO:40, 1349381T6 (LATRTUT02) nucleotide 1769-1583 of SEQ ID NO:40, 2683322CT1 (SINIUCT01), 2683322H1 (SINIUCT01) nucleotide 496-598 of SEQ ID NO:40
10	41	2684552	LUNGNOT23	267589R6 (HNT2NOT01) nucleotide 22-659 of SEQ ID NO:41, 1384315F1 (BRAITUT08) nucleotide 987-1588 of SEQ ID NO:41, 1622931F6 (BRAITUT13) nucleotide 935-1380 of SEQ ID NO:41, 1707466F6 (DUODNOT02) nucleotide 1379-1825 of SEQ ID NO:41, 2470827H1 (THP1NOT03) nucleotide 1-232 of SEQ ID NO:41, 2684552H1 (LUNGNOT23) nucleotide 16-273 of SEQ ID NO:41, SASA00264F1 nucleotide 299-614 of SEQ ID NO:41

Table 1 (cont.)

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
11	42	2830310	TLYMNOT03	2594867T6 (OVARTUT02) nucleotide 1933-1516 of SEQ ID NO:42, 2830310CT1 (TLYMNOT03), 2830310H1 (TLYMNOT03) nucleotide 674-873 of SEQ ID NO:42
12	43	2963346	SCORNOT04	775148F1 (COLNNOT05) nucleotide 2698-2123 of SEQ ID NO:43, 1427027F1 (SINTBST01) nucleotide 4-456 of SEQ ID NO:43, 1889805F6 (BLADTUT07) nucleotide 2237-2709 of SEQ ID NO:43, 2963346CT1 (SCORNOT04), 2963346H1 (SCORNOT04) nucleotide 2120-2417 of SEQ ID NO:43
13	44	2994234	KIDNFET02	999410H1 (KIDNTUT01) nucleotide 1-139 of SEQ ID NO:44, 2272381R6 (PROSNON01) nucleotide 998-1458 of SEQ ID NO:44, 2957657F6 (KIDNFET01) nucleotide 326-930 of SEQ ID NO:44, 2994234F6 (KIDNFET02) nucleotide 582-1098 of SEQ ID NO:44, 2994234H1 (KIDNFET02) nucleotide 583-862 of SEQ ID NO:44
14	45	4115958	UTRSTUT07	1285928T6 (COLNNOT16) nucleotide 763-1 of SEQ ID NO:45, 1502647F1 (BRAITUT07) nucleotide 1878-2403 of SEQ ID NO:45, 1816330F6 (PROSNOT20) nucleotide 1272-1881 of SEQ ID NO:45, 1816330T6 (PROSNOT20) nucleotide 2733-2338 of SEQ ID NO:45, 2778671T6 (OVARTUT03) nucleotide 2402-1825 of SEQ ID NO:45, 2907759H1 (THYMNOT05) nucleotide 883-1135 of SEQ ID NO:45, 2998886H1 (OVARTUT07) nucleotide 714-966 of SEQ ID NO:45, 3037339H1 (BRSTNOT16) nucleotide 1080-1361 of SEQ ID NO:45, 4115958H1 (UTRSTUT07) nucleotide 1761-1927 of SEQ ID NO:45
15	46	779255	MYOMNOT01	779255H1 (MYOMNOT01), 779559X11 (MYOMNOT01), 874628R1 (LUNGAST01), 1229020R6 (BRAITUT01), 2186473T6 (PROSNOT26)
16	47	1303605	PLACNOT02	287928F1 (EOSIHET02), 617771R6 (PGANNOT01), 1303605H1 (PLACNOT02), 1303605X15 (PLACNOT02), 1303605X19 (PLACNOT02), 1551004R6 (PROSNOT06), 1580268F6 (DUODNOT01)

Table 1 (cont.)

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
17	48	1611167	COLNTUT06	001705X2 (U937NOT01), 1338850F6 (COLNTUT03), 1338850T6 (COLNTUT03), 1611167F6 (COLNTUT06), 1611167H1 (COLNTUT06), 2073827T6 (ISLTNOT01), 2113049H1 (BRAITUT03), 2995424H1 (OVARTUT07), 3728823F6 (SMCCNON03)
18	49	1907472	CONNTUT01	884279T1 (PANCNOT05), 1336640F6 (COLNNOT13), 1687941F6 (PROSTUT10), 1871741F6 (LEUKNOT02), 1907472F6 (CONNTUT01), 1907472H1 (CONNTUT01), 2137210F6 (ENDCNOT01), 2653156H1 (THYMNOT04), 2726334F6 (OVARTUT05), 3520477T6 (LUNGNON03), 4176848H1 (BRAINOT22)
19	50	1985458	LUNGAST01	1223005R1 (COLNTUT02), 1271110F1 (TESTTUT02), 1985458H1 (LUNGAST01), 2603571H1 (LUNGTUT07)
20	51	2726431	OVARTUT05	2719325T6 (THYRNOT09), 2726431H1 (OVARTUT05), SAEA01318R1, SAEA10035P1, SAEA02442R1, SAEA01598R1, SAEA02361R1
21	52	2743828	BRSTTUT14	2211446F6 (SINTFET03), 2743828F6 (BRSTTUT14), 2743828H1 (BRSTTUT14), 2885037H1 (SINJNOT02), 4637959H1 (MYEPTXT01)
22	53	2998209	OVARTUT07	1307477F6 (COLNFET02), 1440536F1 (THYRNOT03), 1449582F6 (PLACNOT02), 1503422T6 (BRAITUT07), 1904631H1 (OVARNOT07), 2587988H1 (BRAITUT22), 2887687H1 (SINJNOT02), 2907651H1 (THYMNOT05), 2998209F6 (OVARTUT07), 2998209H1 (OVARTUT07), 3113687F6 (BRSTNOT17), 3113687T6 (BRSTNOT17)
23	54	3340296	SPLNNOT10	954226X13 (KIDNNOT05), 1307821R1 (COLNFET02), 1851913F6 (LUNGFET03), 2890089H1 (LUNGFET04), 3340296H1 (SPLNNOT10)
24	55	3536740	KIDNNOT25	3536740F6 (KIDNNOT25), 3536740H1 (KIDNNOT25), 3764761H1 (PROSTUT13)

Table 1 (cont.)

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
25	56	082155	HUVESTB01	082155H1 (HUVESTB01), 292288F1 (TMLR3DT01), 522294R6 (MMLR2DT01), 1376713H1 (LUNGNOT10), 1668753F6 (BMARNOT03), 1874447T6 (LEUKNOT02), 1880012F6 (LEUKNOT03), 2688343H1 and 2693743T6 (LUNGNOT23), 3084261H1 (BRAIFET01)
26	57	095477	PITUNOT01	1350375F6 (LATRTUT02), 1818242H1 (PROSNOT20), 2579350F6 (KIDNTUT13), 3453839H1 and 3535405H1 (KIDNNOT25), 095477H1, 095477R6, and 095477X3 (PITUNOT01), 129825R6 (TESTNOT01), 276986T6 (TESTNOT03), 1479477F1 (CORPNOT02), 1624466F6 and 1624466T6 (BRAITUT13), 1626694F6 (COLNPOT01), 1917535H1 (PROSNOT06), 2133306F6 (ENDCNOT01), 2474044F6 (THP1NOT03), 2844721H1 (DRGLNOT01), 3099615F6 (PTHYNOT03), 3880403H1 (SPLNNOT11)
27	58	1399169	BRAITUT08	412648R1 (BRSTNOT01), 1399169H1 (BRAITUT08), 1428917F7 (SINTBST01), 1931149F6 (COLNTUT03), 2153060H1 (BRAINOT09), 2752020H1 (THP1AZS08), 2990293X311F1 and 2992892F6 (KIDNFET02), 3345146F6 (SPLNNOT09), 3799739H1 (SPLNNOT12), 3937158H1 (SKINBIT01)
28	59	1442069	THYRNOT03	1442069H1 and 1442069R1 (THYRNOT03), SBJA03794F1, SBJA00434F1, SBJA01969F1
29	60	1596668	BRAINOT14	1596668H1 (BRAINOT14), 1793316T6 (PROSTUT05), SAGA03432F1, SAGA02842F1, SAGA00436R1

Table 1 (cont.)

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
30	61	1977214	PANCTUT02	452406R6 (TLYMNOT02), 1259390F6 (MENITUT03), 1444230R1 (THYRNOT03), 1695371F6 (COLNNOT23), 1903126F6 (OVARNOT07), 1977214H1 (PANCTUT02), 2058371R6 (OVARNOT03), 2357444H1 (LUNGNOT20), 2706531H1 (PONSAZT01), 2850271H1 (BRSTTUT13), 2917485H1 (THYMFET03), SAEA00128F1, SAEA00912R1, SAEA03455F1
31	62	2181282	SININOT01	2181282F6 and 2181282H1 (SININOT01), 2538117F6 (BONRTUT01), 2970580F6 (HEAONOT02), 3529310T6 (BLADNOT09)

Table 2

Polypeptide Seq ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequence	Identification	Analytical Methods
1	379	S258 S276 S32 S102 S128 S141 S260 S276 T342 S350 S12 T67 S91 S121 S166 S249 S282 T314 T328 T367 S363 Y58	N126 N242	G60-T67	Nucleolar zinc finger protein; cell growth regulator; nucleotide-binding	MOTIFS BLAST
2	136	T4 T86 S107 T90 S129	N127		Novel open reading frame	
3	230	S126 S35 T170		K131-E188 I119-R124 I163-K186	Homeodomain protein	BLIMPS (BLOCKS) BLIMPS (PRINTS) PFAM PROFILESCAN BLAST
4	131	S70 T5 S14 S123	N122		Krueppel-related zinc finger protein	BLAST
5	411	T33 S81 T102 S141 S156 T339 S356 T357 S363 S370 T381 T389 S71 S123 S181 T187 T215	N121	P39-E112 P60-H128 L110-I118	HMG protein, chromodomain signature	BLIMPS (BLOCKS) BLIMPS (PRINTS) PFAM

Table 2 (Cont.)

Polypeptide Seq ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequence	Identification	Analytical Methods
6	226	T9 T45 T80 S167 T204 S212 S24 T45 S122 T125 Y32 Y62	N92		Histone 2A-related protein with leucine zipper	BLAST
7	183	S29 S22 S109 T46 S170			Homeodomain protein	BLAST
8	317	S50 T189 S221 T232 S291 S67 S156 S201 T290 Y55	N208	V28-A248 comprises 5 signatures; I110-A128	Transcription factor	BLIMPS (BLOCKS) MOTIFS BLAST
9	479	T48 S93 S154 T220 T276 S315 S316 S90 T142 S159 T178 S257 T306 T342 T430	N420	R98-G329 S347-G434	RNA helicase, DEAD-box subfamily	BLIMPS (BLOCKS) PFAM BLAST
10	582	T26 T41 S49 T130 S256 S278 T414 T437 S487 T63 S327 S534 S540 T562 Y108	N24 N196 N203 N457	C42-C83 C39-H59	Cysteine rich, zinc finger protein	PFAM
11	327	S225 S16 S67 S68 S144 T193 Y185	N45	I154-R177 K205-V252 C290-K324	Glucose-repressible alcohol dehydrogenase transcriptional effector, AP endonuclease motif	BLIMPS (BLOCKS) BLAST

Table 2 (Cont.)

Polypeptide Seq ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequence	Identification	Analytical Methods
12	502	S15 S41 S94 T173 S309 S328 S389 S3 S249 T280 S301 S391	N295 N366		Novel open reading frame	
13	375	S41 T75 S96 T142 S194 T240 T277 S355 S11 S16 T37 T114 Y53	N338 N353	L221-T240	Transcription termination factor with leucine zippers, helix-turn-helix protein	BLIMPS (PRINTS) BLAST
14	341	S14 S15 T88 T136 T179 S195 S256 S262 S58 S68 T116 T144 S169 S199 S252 T276 S295 S335 S336	N61 N94 N134 N172 N250		Novel open reading frame	
15	269	T5 S84 T118	N20 N38	C11-H31, C39-H59, C67-H89, C97-H117, C127-H147	zinc finger (C2H2) protein	MOTIFS PFAM BLOCKS
16	264	S138 S10 T137 S180 S242 S52 S76	N178 N215 N260	C206-C240	zinc finger (C3HC4) protein; apoptosis inhibitor	PFAM BLOCKS BLAST

Table 2 (Cont.)

Polypeptide Seq ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequence	Identification	Analytical Methods
17	605	S141 S335 S353 T15 T24 S217 S34 S90 T93 S112 T148 S158 S169 S195 S228 S297 S308 S349 T362 S368 T502 T558 S563 Y214	N225 N235 N506	C216-H236, C244-H264, C272-H292, C300-H320, C356-H376, C384-H404, C412-H432, C440-H460, C468-H488, C496-H516, C524-H544, C552-H572, C580-H600	zinc finger (C2H2) protein	MOTIFS PFAM BLOCKS PRINTS
18	757	S28 T46 T59 T69 T156 T332 S338 S367 S374 S436 T494 S574 T650 S713 T724 T190 T205 T257 S307 S421 S480 S483 S536 T658 Y143 Y199 Y750	N75 N187	C104-K115 H71-L122	cysteine rich, metal binding; signal transduction asssociated	BLOCKS PFAM
19	154	S7 T71 S94 T143 T64 T137	N92	H45-P115, T127-C136 A114-V154	cysteine rich, metal binding; zinc finger (C2H2) protein	BLOCKS PRINTS

Table 2 (Cont.)

Polypeptide Seq ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequence	Identification	Analytical Methods
20	587	S117 T234 T462 T14 S51 S477 S566 S571 S141 T229 T257 S313 S341 S369 S397 S537 T547 S568 Y107 Y254 Y282 Y439 Y467	N62 N533	C133-H153, C161-H181, C189-H209, C217-H237, C245-H265, C273-H293, C301-H321, C329-H349, C357-H377, C385-H405, C413-H433, C441-H461, C469-H489, C497-H517, C525-H545	zinc finger (C2H2) protein	MOTIFS PFAM BLOCKS PRINTS BLAST
21	346	T176 S185 T36 T71 S137 S188 S206 S231 S184 T284 Y256	N292	E283-L303 C277-D305	cysteine rich, metal binding; zinc finger (C2H2) protein; similar to D. melanogaster trithorax protein	BLOCKS PRINTS BLAST

22	481	S14 S60 T69 S80 T166 T199 T212 S242 S270 T381 S406 S468 S470 S320 S395 S440 S441 Y146 Y261	N192 N210		similar to M. musculus A20 (zinc finger protein, inhibitor of apoptosis)	BLAST

Table 2 (Cont.)

Polypeptide Seq ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential glycosylation sites	Signature Sequence	Identification	Analytical Methods
23	179	S5 S7 S27 T33 S37 S42 S116 T64 T120	N35 N149	S96-T151 Q80-R132	Myc-type HLH protein	PROFILESCAN MOTIFS PFAM BLOCKS
24	254	S186 S94 S96 S134 T189 S41 T102		G122-S186 P108-V167 L168-Q202	Homeobox proteins; transcriptional repressor; HMG protein	PROFILESCAN MOTIFS PFAM BLOCKS PRINTS
25	498	T242 S383 S384 S50 S87 S244 S425 S4 T40 T130 T263 S271 T325 T330 S352 S362	N393 N458	I9-N67 C30-I39	zinc finger protein (gpStaf50)	BLAST, BLOCKS PFAM, MOTIFS ProfileScan

Table 2 (Cont.)

Polypeptide Seq ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequence	Identification	Analytical Methods
26	1299	T85 S139 S160 S199 S225 S277 S284 S335 S340 T382 S392 T412 S424 T480 S505 S684 S690 T702 S708 S945 S1001 S1067 S1166 S1253 T1295 T105 S122 S225 S308 T351 T359 S377 T405 T461 T665 T998 T1092 S1108 S1260 S1285 Y43 Y136 Y268 Y168	N91 N934 N991 N1031 N1090 N1098 N1235 N1246 N1282		glutamine rich protein	BLAST MOTIFS
27	951	T101 T163 S279 S20 T91 S136 S187 T206 T207 S398 S407 S439 T480 T488 S517 S593 T605 S683 S714 S729 S738 T755 S841 S862 S915 S930 T5 S20 S60 T77 T147 S284 S439 T552 S683 T829 S899 S902 S908 T909 S915 Y167	N289 N301 N748 N860	P808-Q880	bromodomain protein	BLAST, BLOCKS PFAM, MOTIFS PRINTS, ProfileScan

Table 2 (Cont.)

Polypeptide Seq ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequence	Identification	Analytical Methods
28	282	S20 S40 T205 T47 T66 S93 S41 S42 T109		R13-R43 N9-E89	HMG-14 protein	BLOCKS, MOTIFS
29	186	T72 S73 T169 S181 S63 S144 S181		T14-T72 V103-N163	Transcription initiation factor (TFIID)	BLAST, MOTIFS BLOCKS PRINTS, PFAM
30	917	S32 S86 S90 S395 S492 S497 S808 T118 S132 S143 S159 S195 S235 S284 T307 S331 S350 S373 S383 S384 S455 T456 S505 S627 T28 T85 S89 T224 S283 T391 T416 S492 S507 S547 T560 S576 S627 S803 S874 T914 Y702 Y736	N53 N141	M1-S24 F408-V432	RNA-binding protein (RNP-1)	SPScan, PFAM BLOCKS, MOTIFS
31	392	S193 S222 T253 S279 S280 S388 T39 S52 T94 S100 S179 T314	N35 N49 N191	M1-S22 C73-C101 C132-H154 C342-H363 P98-P111 L114-G123	Zinc finger protein	BLAST, MOTIFS BLOCKS, PRINTS SPScan, PFAM

Table 3

Nucleotide SEQ ID NO:	Tissue Expression (Fraction of Total)	Disease Class (Fraction of Total)	Vector
32	Reproductive (0.381) Hematopoietic/Immune (0.190) Cardiovascular (0.167)	Cancer (0.429) Inflammation (0.310) Fetal (0.286)	pSPORT1
33	Reproductive (0.286) Nervous (0.224) Hematopoietic/Immune (0.122)	Cancer (0.490) Inflammation (0.327) Fetal (0.122)	pSPORT1
34	Reproductive (0.500) Cardiovascular (0.167) Urologic (0.167)	Cancer (0.667) Trauma (0.333)	pSPORT1
35	Reproductive (0.357) Hematopoietic/Immune (0.214) Developmental (0.143)	Cancer (0.286) Inflammation (0.286) Fetal (0.214)	pINCY
36	Reproductive (0.265) Gastrointestinal (0.126) Hematopoietic/Immune (0.119)	Cancer (0.490) Fetal (0.219) Inflammation (0.219)	pINCY
37	Developmental (0.429) Cardiovascular (0.143) Gastrointestinal (0.143)	Fetal (0.571) Cancer (0.429) Inflammation (0.143)	pINCY
38	Reproductive (0.304) Nervous (0.203) Gastrointestinal (0.174)	Cancer (0.522) Inflammation (0.261) Fetal (0.116)	pINCY
39	Reproductive (0.286) Cardiovascular (0.143) Hematopoietic/Immune (0.143)	Cancer (0.714) Inflammation (0.143)	pBLUESCRIPT
40	Reproductive (0.239) Gastrointestinal (0.174) Cardiovascular (0.130)	Cancer (0.457) Inflammation (0.326) Fetal (0.196)	pINCY
41	Nervous (0.256) Cardiovascular (0.186) Gastrointestinal (0.140)	Cancer (0.535) Inflammation (0.233) Fetal (0.163)	pINCY
42	Cardiovascular (0.200) Reproductive (0.200) Hematopoietic/Immune (0.150)	Cancer (0.350) Inflammation (0.250) Fetal (0.150)	pINCY

Table 3 (Cont.)

Nucleotide SEQ ID NO:	Tissue Expression (Fraction of Total)	Disease Class (Fraction of Total)	Vector
43	Reproductive (0.220) Nervous (0.165) Cardiovascular (0.143)	Cancer (0.440) Inflammation (0.253) Fetal (0.198)	pINCY
44	Reproductive (0.242) Nervous (0.212) Developmental (0.121)	Cancer (0.424) Inflammation (0.242) Trauma (0.182)	pINCY
45	Reproductive (0.277) Nervous (0.191) Gastrointestinal (0.149)	Cancer (0.553) Inflammation (0.234) Trauma (0.149)	pINCY
46	Reproductive (0.273) Nervous (0.227) Hematopoietic/Immune (0.205)	Cancer (0.455) Inflammation (0.273) Other (0.114)	pSPORT1
47	Reproductive (0.222) Developmental (0.167) Cardiovascular (0.153)	Cancer (0.403) Inflammation (0.236) Fetal (0.222)	pINCY
48	Hematopoietic/Immune (0.241) Reproductive (0.241) Nervous (0.207)	Cancer (0.517) Inflammation (0.276) Fetal (0.207)	pINCY
49	Reproductive (0.192) Nervous (0.178) Gastrointestinal (0.137)	Cancer (0.411) Fetal (0.192) Inflammation (0.192)	pINCY
50	Reproductive (0.236) Nervous (0.173) Gastrointestinal (0.139)	Cancer (0.529) Fetal (0.197) Inflammation (0.192)	pSPORT1
51	Reproductive (0.333) Developmental (0.167) Gastrointestinal (0.167)	Cancer (0.583) Fetal (0.250) Inflammation (0.167)	pINCY
52	Reproductive (0.333) Cardiovascular (0.167) Dermatologic (0.167)	Fetal (0.417) Cancer (0.250) Trauma (0.250)	pINCY
53	Reproductive (0.264) Nervous (0.226) Gastrointestinal (0.132)	Cancer (0.377) Inflammation (0.283) Fetal (0.189)	pINCY

Table 3 (Cont.)

Nucleotide SEQ ID NO:	Tissue Expression (Fraction of Total)	Disease Class (Fraction of Total)	Vector
54	Reproductive (0.275) Developmental (0.225) Urologic (0.150)	Cancer (0.425) Fetal (0.300) Inflammation (0.175)	pINCY
55	Nervous (0.333) Reproductive (0.333) Urologic (0.333)	Cancer (0.667) Inflammation (0.333)	pINCY
56	Hematopoietic/Immune (0.226) Reproductive (0.204) Nervous (0.183)	Cell proliferative (0.602) Inflammation (0.333)	pBLUESCRIPT
57	Reproductive (0.329) Nervous (0.176) Gastrointestinal (0.094)	Cell proliferative (0.694) Inflammation (0.176)	pBLUESCRIPT
58	Hematopoietic/Immune (0.171) Reproductive (0.171) Gastrointestinal (0.158)	Cell proliferative (0.605) Inflammation (0.276)	pINCY
59	Nervous (0.333) Reproductive (0.200) Cardiovascular (0.133)	Cancer (0.467) Trauma (0.267) Other (0.133)	pINCY
60	Reproductive (0.260) Nervous (0.140) Cardiovascular (0.120)	Cell proliferative (0.760) Inflammation (0.120)	pINCY
61	Reproductive (0.234) Nervous (0.188) Gastrointestinal (0.141)	Cell proliferative (0.633) Inflammation (0.219)	pINCY
62	Nervous (0.417) Cardiovascular (0.250) Gastrointestinal (0.167)	Trauma (0.333) Cancer (0.250) Inflammation (0.167)	pINCY

Table 4

Nucleotide SEQ ID NO:	Clone ID	Library	Library Description
32	591290	UTRSNOT01	UTRSNOT01 library was constructed using RNA isolated from the uterine tissue of a 59-year-old female who died of a myocardial infarction. Patient history included cardiomyopathy, coronary artery disease, myocardial infarctions, hypercholesterolemia, hypotension, and arthritis.
33	815856	OVARTUT01	OVARTUT01 library was constructed using RNA isolated from ovarian tumor tissue removed from a 43-year-old Caucasian female during a bilateral salpingo-oopherectomy. Pathology indicated grade 2 mucinous cystadenocarcinoma of the left ovary and involving the entire ovary. Patient history included mitral valve disorder and viral hepatitis. Family history included atherosclerotic coronary artery disease, pancreatic cancer, stress reaction, cerebrovascular disease, breast cancer, and uterine cancer.
34	996352	KIDNTUT01	KIDNTUT01 library was constructed using RNA isolated from kidney tumor tissue removed from an 8-month-old female during nephroureterectomy. Pathology indicated Wilms' tumor (nephroblastoma) involving the renal parenchyma and a capsular blood vessel. Patient history included heparin anticoagulant therapy.
35	1273778	TESTTUT02	TESTTUT02 library was constructed using RNA isolated from testicular tumor tissue removed from a 31-year-old Caucasian male during unilateral orchiectomy. Pathology indicated embryonal carcinoma forming a largely necrotic mass involving the entire testicle. Rare foci of residual testicle showed intralobular germ cell neoplasia, and tumor was identified at the spermatic cord margin.

Table 4 (Cont.)

Nucleotide SEQ ID NO:	Clone ID	Library	Library Description
36	1509715	LUNGNOT14	LUNGNOT14 library was constructed using RNA isolated from lung tissue removed from the left lower lobe of a 47-year-old Caucasian male during a segmental lung resection. Pathology for the associated tumor tissue indicated a grade 4 adenocarcinoma, and the parenchyma showed calcified granuloma. Patient history included benign hypertension and chronic obstructive pulmonary disease. Family history included benign hypertension, Type II diabetes, and acute myocardial infarction.
37	1676367	BLADNOT05	BLADNOT05 library was constructed using RNA isolated from bladder tissue removed from a 60-year-old Caucasian male during a radical cystectomy, prostatectomy, and vasectomy. Pathology for the associated tumor tissue indicated grade 3 transitional cell carcinoma in the left bladder wall with extension through the muscularis propria into the perivascular fat. Carcinoma in situ was identified in the dome and trigone. The prostate showed adenofibromatous hyperplasia. Family history included Type I diabetes, malignant stomach neoplasm, atherosclerotic coronary artery disease, and acute myocardial infarction.
38	1734119	COLNNOT22	COLNNOT22 library was constructed using RNA isolated from colon tissue removed from a 56-year-old Caucasian female with Crohn's disease during a partial resection of the small intestine. Pathology indicated Crohn's disease of the ileum and ileal-colonic anastomosis. The ileal mucosa showed linear and puncture ulcers. Patient history included obesity, a partial ileal resection, permanent ileostomy, cholecystectomy, and excision of breast lesions. Family history included irritable bowel syndrome and atherosclerosis.
39	1944813	PITUNOT01	PITUNOT01 library was constructed using RNA isolated from the normal pituitary glands of 18 male and female Caucasian donors, 16 to 70 years old, who died from trauma. RNA was obtained from Clontech, CLON 6584-2, lot 35278.

Table 4 (Cont.)

Nucleotide SEQ ID NO:	Clone ID	Library	Library Description
40	2683322	SINIUCT01	SINIUCT01 library was constructed using RNA isolated from ileum tissue removed from a 42-year-old Caucasian male during a total intra-abdominal colectomy and endoscopic jejunostomy. Severely active chronic ulcerative colitis was present. The end of the distal colon was completely ulcerated. The proximal colon and cecum showed mild to moderate active colitis. Family history included benign hypertension, cerebrovascular disease, atherosclerotic coronary artery disease, and Type II diabetes.
41	2684552	LUNGNOT23	LUNGNOT23 library was constructed using RNA isolated from lung tissue removed from the left lobe a 58-year-old Caucasian male during a segmental lung resection. Pathology for the associated tumor tissue indicated that the left lower lobe contained metastatic grade 3 (of 4) osteosarcoma, forming four nodules. The left pleura showed metastatic grade 3 (of 4) osteosarcoma. Patient history included soft tissue cancer, secondary cancer of the lung, prostate cancer, acute duodenal ulcer, and benign hypertension. Family history included prostate cancer, benign hypertension, breast cancer, and acute leukemia.
42	2830310	TLYMNOT03	TLYMNOT03 library was constructed using RNA isolated from resting Th1 cells which were differentiated from umbilical cord CD4 T cells with IL-12 and B7-transfected COS cells.
43	2963346	SCORNOT04	SCORNOT04 library was constructed using RNA isolated from cervical spinal cord tissue removed from a 32-year-old Caucasian male who died from acute pulmonary edema, acute bronchopneumonia, bilateral pleural effusions, pericardial effusion, and malignant lymphoma. Patient history included probable cytomegalovirus infection, hepatic congestion and steatosis, splenomegaly, hemorrhagic cystitis, thyroid hemorrhage, respiratory failure, pneumonia of the left lung, pharyngeal natural killer cell lymphoma, and Bell's palsy.
44	2994234	KIDNFET02	KIDNFET02 library was constructed using RNA isolated from heart tissue removed from a Caucasian male fetus who was still-born with a hypoplastic left heart at 23 weeks' gestation.

Table 4 (Cont.)

Nucleotide SEQ ID NO:	Clone ID	Library	Library Description
45	4115958	UTRSTUT07	UTRSTUT07 library was constructed using RNA isolated from uterine tumor tissue removed from a 41-year-old Caucasian female during total abdominal hysterectomy with removal of a single ovary. Pathology indicated the endometrium was secretory phase, and the cervix showed microglandular hyperplasia. There were multiple (2 subserosal, 13 intramural, 1 submucosal) leiomyomas. Family history included atherosclerotic coronary artery disease, benign hypertension, depression, and Type II diabetes.
46	779255	MYOMNOT01	MYOMNOT01 pSPORT1 Library was constructed using RNA isolated from uterine myometrial tissue removed from a 43-year-old Caucasian female during a vaginal hysterectomy and removal of the fallopian tubes and ovaries. Family history included lung cancer, stroke, type II diabetes, hepatic lesion, chronic liver disease, hyperlipidemia, congenital heart anomaly, and mitral valve prolapse.
47	1303605	PLACNOT02	PLACNOT02 pINCY Library was constructed using RNA isolated from the placental tissue of a Hispanic female fetus who was prematurely delivered at 21 weeks' gestation. Serologies of the mother's blood were positive for cytomegalovirus.
48	1611167	COLNTUT06	COLNTUT06 pINCY Library was constructed using RNA isolated from colon tumor tissue obtained from a 45-year-old Caucasian female during a total colectomy and total abdominal hysterectomy. Pathology indicated invasive grade 2 colonic adenocarcinoma forming a cecal mass. Patient history included benign neoplasms of the rectum and anus, multiple sclerosis, mitral valve disorder, and a prior polypectomy. Family history included type I diabetes, cerebrovascular disease, atherosclerotic coronary artery disease, malignant skin neoplasm, hypertension, and malignant neoplasm of the colon.
49	1907472	CONNTUT01	CONNTUT01 pINCY Library was constructed using RNA isolated from a soft tissue tumor removed from the clival area of the skull of a 30-year-old Caucasian female. Pathology indicated chondroid chordoma with neoplastic cells reactive for keratin.

Table 4 (Cont.)

Nucleotide SEQ ID NO:	Clone ID	Library	Library Description
50	1985458	LUNGAST01	LUNGAST01 pSPORT1 Library was constructed using RNA isolated from the lung tissue of a 17-year-old Caucasian male who died from head trauma. Patient history included asthma.
51	2726431	OVARTUT05	OVARTUT05 pINCY Library was constructed using RNA isolated from ovarian tumor tissue removed from a 62-year-old Caucasian female during a total abdominal hysterectomy, removal of the fallopian tubes and ovaries, exploratory laparotomy, regional lymph node excision, and dilation and curettage. Pathology indicated grade 4 endometrioid carcinoma with extensive squamous differentiation, forming a solid mass in the right ovary. The cervix showed mild chronic cervicitis, and focal endometriosis was observed in the posterior uterine serosa. Curettings indicated weakly proliferative endometrium with excessive stromal breakdown in the uterus, and a prior cervical biopsy indicated mild chronic cervicitis with a prominent nabothian cyst in the cervix. Patient history included longitudinal deficiency of the radioulna, osteoarthritis, thrombophlebitis, and abnormal blood chemistries. Family history included atherosclerotic coronary artery disease, pulmonary embolism, and cerebrovascular disease.
52	2743828	BRSTTUT14	BRSTTUT14 pINCY Library was constructed using RNA isolated from breast tumor tissue removed from a 62-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology indicated an invasive grade 3 (of 4), nuclear grade 3 (of 3) adenocarcinoma, ductal type. Ductal carcinoma in situ, comedo type, comprised 60% of the tumor mass. Metastatic adenocarcinoma was identified in one of 14 axillary lymph nodes. Tumor cells were strongly positive for estrogen receptors and weakly positive for progesterone receptors. Patient history included benign colon neoplasm, hyperlipidemia, cardiac dysrhythmia, and obesity. Family history included atherosclerotic coronary artery disease, myocardial infarction, colon cancer, ovarian cancer, lung cancer, and cerebrovascular disease.

Table 4 (Cont.)

Nucleotide SEQ ID NO:	Clone ID	Library	Library Description
53	2998209	OVARTUT07	OVARTUT07 pINCY Library was constructed using RNA isolated from right ovarian tumor tissue removed from a 58-year-old Caucasian female during bilateral salpingo-oophorectomy, regional lymph node excision, destruction of peritoneal tissue, cystocele repair, and skin repair. Pathology indicated grade 3 adenocarcinoma, serous type, forming a mass and entirely replacing the right ovary. The left pelvic sidewall revealed a microscopic focus of metastatic adenocarcinoma. Patient history included hyperlipidemia, thrombophlebitis, and carcinoma in situ of the cervix uteri. Family history included cerebrovascular disease, breast cancer, hyperlipidemia, atherosclerotic coronary artery disease, and heart failure.
54	3340296	SPLNNOT10	SPLNNOT10 pINCY Library was constructed using RNA isolated from spleen tissue removed from a 59-year-old Caucasian male during a total splenectomy and exploratory laparotomy. Pathology indicated splenomegaly with congestion. The lymph nodes showed reactive follicular hyperplasia. The liver showed mild, nonspecific steatosis. A portion of the spleen contained abundant CD3- and CD5-positive T-lymphocytes and CD19- and CD20-positive B-lymphocytes that stained immunocytochemically for kappa and lambda immunoglobulin light chains. Patient history included poliovirus infection. Family history included myocardial infarction, arteriosclerotic cardiovascular disease, primary tuberculous infection, cerebrovascular disease and lymphoma.
55	3536740	KIDNNOT25	KIDNNOT25 pINCY Library was constructed using RNA isolated from kidney tissue removed from the left lower kidney pole of a 42-year-old Caucasian female during nephroureterectomy. Pathology indicated slight hydronephrosis and nephrolithiasis. Patient history included calculus of the kidney.
56	082155	HUVESTB01	The HUVESTB01 library was constructed using RNA isolated from shear-stressed HUV-EC-C (ATCC CRL 1730) cells. Before RNA isolation, the cells were subjected to a shear stress of 10 dynes/cm.

Table 4 (Cont.)

SEQ ID NO:	Clone ID	Library	Library Comment
57	095477	PITUNOT01	The PITUNOT01 library was constructed using RNA obtained from Clontech (CLON 6584-2, lot 35278). The RNA was isolated from the pituitary glands removed from a pool of 18 male and female Caucasian donors, 16 to 70 years old, who died from trauma. RNA was isolated by a modified GuSCN method, followed by two rounds of polyA RNA selection on oligo(dT)-cellulose columns.
58	1399169	BRAITUT08	The BRAITUT08 library was constructed using RNA isolated from brain tumor tissue removed from the left frontal lobe of a 47-year-old Caucasian male during excision of cerebral meningeal tissue. Pathology indicated grade 4 fibrillary astrocytoma with focal tumoral radionecrosis. Patient history included cerebrovascular disease, deficiency anemia, hyperlipidemia, epilepsy, and tobacco use. Family history included cerebrovascular disease and a malignant prostate neoplasm.
59	1442069	THYRNOT03	The THYRNOT03 library was constructed using RNA isolated from thyroid tissue removed from the left thyroid of a 28-year-old Caucasian female during a complete thyroidectomy. Pathology indicated a small nodule of adenomatous hyperplasia present in the left thyroid. Pathology for the associated tumor tissue indicated dominant follicular adenoma forming a well-encapsulated mass in the left thyroid.
60	1596668	BRAINOT14	The BRAINOT14 library was constructed using RNA isolated from brain tissue removed from the left frontal lobe of a 40-year-old Caucasian female during excision of a cerebral meningeal lesion. Pathology for the associated tumor tissue indicated grade 4 gemistocytic astrocytoma.

Table 4 (Cont.)

SEQ ID NO:	Clone ID	Library	Library Comment
61	1977214	PANCTUT02	The PANCTUT02library was constructed using RNA isolated from pancreatic tumor tissue removed from a 45-year-old Caucasian female during radical pancreaticoduodenectomy. Pathology indicated a grade 4 anaplastic carcinoma. Family history included benign hypertension, hyperlipidemia and atherosclerotic coronary artery disease.
62	2181282	SININOT01	The SININOT01 library was constructed using RNA isolated from ileum tissue obtained from the small intestine of a 4-year-old Caucasian female who died from a closed head injury. Patient history included jaundice. Previous surgeries included a double hernia repair.

Table 5

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25: 3389-3402.	*ESTs*: Probability value= 1.0E-8 or less *Full Length sequences*: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises as least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad Sci. 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183: 63-98; and Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489.	*ESTs*: fasta E value=1.06E-6 *Assembled ESTs*: fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less *Full Length sequences*: fastx score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS and PRINTS databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S and J.G. Henikoff, Nucl. Acid Res., 19:6565-72, 1991. J.G. Henikoff and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37: 417-424.	Score=1000 or greater; Ratio of Score/Strength = 0.75 or larger; and Probability value= 1.0E-3 or less
PFAM	A Hidden Markov Models-based application useful for protein family search.	Krogh, A. et al. (1994) J. Mol. Biol., 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322.	Score=10-50 bits, depending on individual protein families

Table 5 cont.

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25: 217-221.	Score= 4.0 or greater
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M. S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12: 431-439.	Score=5 or greater
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch et al. supra; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

What is claimed is:

1. A substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, and fragments thereof.

2. A substantially purified variant having at least 90% amino acid sequence identity to the amino acid sequence of claim 1.

3. An isolated and purified polynucleotide encoding the polypeptide of claim 1.

4. An isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide of claim 3.

5. An isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide of claim 3.

6. An isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide of claim 3.

7. A method for detecting a polynucleotide, the method comprising the steps of:

(a) hybridizing the polynucleotide of claim 6 to at least one nucleic acid in a sample, thereby forming a hybridization complex; and

(b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of the polynucleotide in the sample.

8. The method of claim 7 further comprising amplifying the polynucleotide prior to hybridization.

9. An isolated and purified polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62, and fragments thereof.

10. An isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide of claim 9.

11. An isolated and purified polynucleotide having a sequence which is

complementary to the polynucleotide of claim 9.

12. An expression vector comprising at least a fragment of the polynucleotide of claim 3.

13. A host cell comprising the expression vector of claim 12.

14. A method for producing a polypeptide, the method comprising the steps of:

a) culturing the host cell of claim 13 under conditions suitable for the expression of the polypeptide; and

b) recovering the polypeptide from the host cell culture.

15. A pharmaceutical composition comprising the polypeptide of claim 1 in conjunction with a suitable pharmaceutical carrier.

16. A purified antibody which specifically binds to the polypeptide of claim 1.

17. A purified agonist of the polypeptide of claim 1.

18. A purified antagonist of the polypeptide of claim 1.

19. A method for treating or preventing a disorder associated with decreased expression or activity of PRGE, the method comprising administering to a subject in need of such treatment an effective amount of the pharmaceutical composition of claim 15.

20. A method for treating or preventing a disorder associated with increased expression or activity of PRGE, the method comprising administering to a subject in need of such treatment an effective amount of the antagonist of claim 18.

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/12, C07K 14/47, 16/18, A61K 38/17

A2

(11) International Publication Number: WO 99/64596

(43) International Publication Date: 16 December 1999 (16.12.99)

(21) International Application Number: PCT/US99/13281

(22) International Filing Date: 11 June 1999 (11.06.99)

(30) Priority Data:

60/089,029	12 June 1998 (12.06.98)	US
60/094,575	29 July 1998 (29.07.98)	US
60/104,624	14 October 1998 (14.10.98)	US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications

US	60/094,575 (CIP)
Filed on	29 July 1998 (29.07.98)
US	60/104,624 (CIP)
Filed on	14 October 1998 (14.10.98)
US	60/089,029 (CIP)
Filed on	12 June 1998 (12.06.98)

(71) Applicant *(for all designated States except US)*: INCYTE PHARMACEUTICALS, INC. [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: LAL, Preeti [IN/US]; 2382 Lass Drive, Santa Clara, CA 95054 (US). YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). TANG, Y., Tom [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). HILLMAN, Jennifer, L. [US/US]; 230 Monroe Drive #12, Mountain View, CA 94040 (US). BANDMAN, Olga [US/US]; 366 Anna Avenue, Mountain View, CA 94043 (US). CORLEY, Neil, C. [US/US]; 1240 Dale Avenue #30, Moutain View, CA 94040 (US). GUEGLER, Karl, J. [CH/US]; 1048 Oakland Avenue, Menlo Park, CA 94025 (US). GORGONE, Gina, A. [US/US]; 1253 Pinecrest Drive, Boulder Creek, CA 95005 (US). BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). PATTERSON, Chandra [US/US]; 490 Sherwood Way #1, Menlo Park, CA 94025 (US). LU, Dyung, Aina, M. [US/US]; 55 Park Belmont Place, San Jose, CA 95136 (US).

(74) Agents: BILLINGS, Lucy, J. et al.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: PROTEINS REGULATING GENE EXPRESSION

(57) Abstract

The invention provides human proteins regulating gene expression (PRGE) and polynucleotides which identify and encode PRGE. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating or preventing disorders associated with expression of PRGE.

Docket No.: PF-0539 USN

DECLARATION AND POWER OF ATTORNEY FOR UNITED STATES PATENT APPLICATION

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name, and

I believe that I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if more than one name is listed below) of the subject matter which is claimed and for which a United States patent is sought on the invention entitled

MOLECULES REGULATING GENE EXPRESSION

the specification of which:

/___/ is attached hereto.

/ X __/ was filed on November 28, 2000 as application Serial No. _____ and if this box contains an X /___/, was amended on _____.

/ X _/ was filed as Patent Cooperation Treaty international application No. PCT/US99/13281 on June 11, 1999 if this box contains an X /_/, was amended on under Patent Cooperation Treaty Article 19 on _____ 2001, and if this box contains an X /_/, was amended on _____.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge my duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, §1.56(a).

I hereby claim the benefit under Title 35, United States Code, §119 or §365(a)-(b) of any foreign application(s) for patent or inventor's certificate indicated below and of any Patent Cooperation Treaty international applications(s) designating at least one country other than the United States indicated below and have also identified below any foreign application(s) for patent or inventor's certificate and Patent Cooperation Treaty international application(s) designating at least one country other than the United States for the same subject matter and having a filing date before that of the application for said subject matter the priority of which is claimed:

Docket No.: PF-0539 USN

Country	Number	Filing Date	Priority Claimed
_____	_____	_____	/_/ Yes /_/ No
_____	_____	_____	/_/ Yes /_/ No

I hereby claim the benefit under Title 35, United States Code, §119(e) of any United States provisional application(s) listed below.

Application Serial No.	Filed	Status (Pending, Abandoned, Patented)
60/089,029	June 12, 1998	Expired
60/094,575	July 29, 1998	Expired
60/104,624	October 14, 1998	Expired

I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in said prior application(s) in the manner required by the first paragraph of Title 35, United States Code §112, I acknowledge my duty to disclose material information as defined in Title 37 Code of Federal Regulations, §1.56(a) which occurred between the filing date(s) of the prior application(s) and the national or Patent Cooperation Treaty international filing date of this application:

Application Serial No.	Filed	Status (Pending, Abandoned, Patented)

I hereby appoint the following:

Lucy J. Billings	Reg. No. 36,749
Michael C. Cerrone	Reg. No. 39,132
Diana Hamlet-Cox	Reg. No. 33,302
Richard C. Ekstrom	Reg. No. 37,027
Barrie D. Greene	Reg. No. 46,740
Matthew R. Kaser	Reg. No. 44,817
Lynn E. Murry	Reg. No. 42,918
Shirley A. Recipon	Reg. No. 47,016
Susan K. Sather	Reg. No. 44,316
Michelle M. Stempien	Reg. No. 41,327
David G. Streeter	Reg. No. 43,168
Stephen Todd	Reg. No. 47,139
Christopher Turner	Reg. No. 45,167
P. Ben Wang	Reg. No. 41,420

respectively and individually, as my patent attorneys and/or agents, with full power of substitution and revocation, to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith. Please address all communications to:

Docket No.: PF-0539 USN

LEGAL DEPARTMENT
INCYTE GENOMICS, INC.
3160 PORTER DRIVE, PALO ALTO, CA 94304

TEL: 650-855-0555 FAX: 650-849-8886 or 650-845-4166

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

1-CO

Sole Inventor or First Joint Inventor:

Full name: **PREETI LAL**

Signature: Preeti Lal

Date: 17th January ,2001

Citizenship: India

Residence: Santa Clara, California CA

P.O. Address: P.O. Box 5142
Santa Clara, California 95056

Docket No.: PF-0539 USN

Second Joint Inventor:

2-00

Full name: **HENRY YUE**

Signature:

Date: January 26 ,2001

Citizenship: United States of America

Residence: Sunnyvale, California CA

P.O. Address: 826 Lois Avenue
Sunnyvale, California 94087

Third Joint Inventor:

3-00

Full name: **Y. TOM TANG**

Signature:

Date: February 27 ,2001

Citizenship: ~~People's Republic of China~~ USA
U.S.A. 2/27/2001

Residence: San Jose, California CA

P.O. Address: 4230 Ranwick Court
San Jose, California 95118

Fourth Joint Inventor:

4-00

Full name: **JENNIFER L. HILLMAN**

Signature:

Date: February 16 ,2001

Citizenship: United States of America

Residence: Mountain View, California CA

P.O. Address: 230 Monroe Drive, #17
Mountain View, California 94040

Docket No.: PF-0539 USN

Fifth Joint Inventor:

Full name: 5-00 **OLGA BANDMAN**

Signature: Olga Bandman

Date: 16 February ,2001

Citizenship: United States of America

Residence: Mountain View, California CA

P.O. Address: 366 Anna Avenue
Mountain View, California 94043

Sixth Joint Inventor:

Full name: 6-00 **NEIL C. CORLEY**

Signature:

Date: JANUARY 31 ,2001

Citizenship: United States of America

Residence: Castro Valley, California CA

P.O. Address: 20426 Crow Creek Road
Castro Valley, California 94552

Seventh Joint Inventor:

Full name: 7-00 **KARL J. GUEGLER**

Signature:

Date: 02/02 ,2001

Citizenship: United States of America

Residence: Menlo Park, California CA

P.O. Address: 1045 Oakland Avenue
Menlo Park, California 94025

Docket No.: PF-0539 USN

Eighth Joint Inventor:

Full name: **GINA A. GORGONE**

Signature:

Date: Jan. 17 ,2001

Citizenship: United States of America

Residence: Boulder Creek, California CA

P.O. Address: 1253 Pinecrest Drive
Boulder Creek, California 95006

Ninth Joint Inventor:

Full name: **MARIAH R. BAUGHN**

Signature:

Date: January 17 ,2001

Citizenship: United States of America

Residence: San Leandro, California CA

P.O. Address: 14244 Santiago Road
San Leandro, California 94577

Tenth Joint Inventor:

Full name: **CHANDRA PATTERSON**

Signature:

Date: 1/17 ,2001

Citizenship: United States of America

Residence: Menlo Park, California CA

P.O. Address: 490 Sherwood Way, #1
Menlo Park, California 94025

Docket No.: PF-0539 USN

Eleventh Joint Inventor:

Full name: **DYUNG AINA M. LU**

Signature:

Date: March 22 ,2001

Citizenship: United States of America

Residence: San Jose, California CA

P.O. Address: 233 Coy Drive
San Jose, California 95123

#9

Docket No.: PF-0539 USN

Certificate of Mailing

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to: Box Sequence, Commissioner for Patents, P.O. Box 2327, Arlington, VA 22202 on November 8, 2002.

By: Joyce Abriam Printed: Joyce Abriam

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:	Lal et al.		
Title:	PROTEINS REGULATING GENE EXPRESSION		
Serial No.:	09/701,674	Filing Date:	To Be Assigned
Examiner:	To Be Assigned	Group Art Unit:	To Be Assigned

United States Patent and Trademark Office
Box Sequence, P.O. Box 2327
Arlington, VA 22202

CERTIFICATE UNDER 37 C.F.R. §3.73(b), REVOCATION OF POWER OF ATTORNEY AND APPOINTMENT OF NEW ATTORNEYS

Sir:

The undersigned has reviewed all the documents in the chain of title of the above-identified patent application and, to the best of undersigned's knowledge and belief, title is in the assignee identified above.

Incyte Genomics, Inc., formerly known as Incyte Pharmaceuticals, Inc., having a principal place of business located at 3160 Porter Drive, Palo Alto, California 94304, certifies that it is the assignee and owner of the entire right, title and interest in, to, and under the invention described and claimed in the above-identified application by virtue of an Assignment recorded at Reel 011970, Frame 0446, hereby revokes all previous powers of attorney and appoints the following patent attorneys/agents:

Lucy J. Billings	**Reg. No. 36,749**	**Gina C. Nellesen**	**Reg. No. 52,062**
Jenny Buchbinder	**Reg. No. 48,588**	**Shirley A. Recipon**	**Reg. No. 47,016**
Michael C. Cerrone	**Reg. No. 39,132**	**Cathleen M. Rocco**	**Reg. No. 46,172**
Diana Hamlet-Cox	**Reg. No. 33,302**	**Susan K. Sather**	**Reg. No. 44,316**
Joel Harris	**Reg. No. 44,743**	**Michelle M. Stempien**	**Reg. No. 41,327**
Richard C. Ekstrom	**Reg. No. 37,027**	**David G. Streeter**	**Reg. No. 43,168**
Barrie D. Greene	**Reg. No. 46,740**	**Sreenivasarao Vepachedu**	**Reg. No. 46,395**
Lori L. Kerber	**Reg. No. 41,113**	**James M. Verna**	**Reg. No. 33,287**
Lynn E. Murry	**Reg. No. 42,918**	**Yu-Mei Eureka Wang**	**Reg. No. 50,510**

Docket No.: PF-0539 USN

Please direct all correspondence to:

Legal Department
Incyte Genomics, Inc.
3160 Porter Drive
Palo Alto, California 94304

and direct all telephone calls and facsimile transmissions to: Diana Hamlet-Cox, Incyte Genomics, Inc., Phone: (650) 845-4639, Fax: (650) 849-8886 or (650) 845-4166.

The undersigned (whose title is supplied below) is empowered to act on behalf of the assignee.

I hereby declare that all statements made herein of my own knowledge are true, and that these statements are made with the knowledge that willful false statements, and the like so made, are punishable by fine or imprisonment, or both, under Section 1001, Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

INCYTE GENOMICS, INC.

Date: November 8, 2002

By: Lee Bendekgey
Lee Bendekgey
EVP, General Counsel/Corporate Secretary

SEQUENCE LISTING

<110> INCYTE PHARMACEUTICALS, INC.
LAL, Preeti
YUE, Henry
TANG, Y. Tom
HILLMAN, Jennifer L.
BANDMAN, Olga
CORLEY, Neil C.
GUEGLER, Karl J.
GORGONE, Gina A.
BAUGHN, Mariah R.
PATTERSON, Chandra
LU, Dyung Aina M.

<120> PROTEINS REGULATING GENE EXPRESSION

<130> PF-0539 PCT

<140> To Be Assigned
<141> Herewith

<150> 60/089,029; 60/094,575; 60/104,624
<151> 1998-06-12; 1998-07-29; 1998-10-14

<160> 62

<170> PERL Program

<210> 1
<211> 379
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 591290

<400> 1

```
Met Val Phe Phe Thr Cys Asn Ala Cys Gly Glu Ser Val Lys Lys
  1               5                  10                  15
Ile Gln Val Glu Lys His Val Ser Val Cys Arg Asn Cys Glu Cys
             20                  25                  30
Leu Ser Cys Ile Asp Cys Gly Lys Asp Phe Trp Gly Asp Asp Tyr
             35                  40                  45
Lys Asn His Val Lys Cys Ile Ser Glu Asp Gln Lys Tyr Gly Gly
             50                  55                  60
Lys Gly Tyr Glu Gly Lys Thr His Lys Gly Asp Ile Lys Gln Gln
             65                  70                  75
Ala Trp Ile Gln Lys Ile Ser Glu Leu Ile Lys Arg Pro Asn Val
             80                  85                  90
Ser Pro Lys Val Arg Glu Leu Leu Glu Gln Ile Ser Ala Phe Asp
             95                 100                 105
Asn Val Pro Arg Lys Lys Ala Lys Phe Gln Asn Trp Met Lys Asn
            110                 115                 120
Ser Leu Lys Val His Asn Glu Ser Ile Leu Asp Gln Val Trp Asn
```

```
                125                 130                 135
Ile Phe Ser Glu Ala Ser Asn Ser Glu Pro Val Asn Lys Glu Gln
                140                 145                 150
Asp Gln Arg Pro Leu His Pro Val Ala Asn Pro His Ala Glu Ile
                155                 160                 165
Ser Thr Lys Val Pro Ala Ser Lys Val Lys Asp Ala Val Glu Gln
                170                 175                 180
Gln Gly Glu Val Lys Lys Asn Lys Arg Glu Lys Lys Glu Glu Arg
                185                 190                 195
Gln Lys Lys Arg Lys Arg Glu Lys Lys Glu Leu Lys Leu Glu Asn
                200                 205                 210
His Gln Glu Asn Ser Arg Asn Gln Lys Pro Lys Lys Arg Lys Lys
                215                 220                 225
Gly Gln Glu Ala Asp Leu Glu Ala Gly Gly Glu Glu Val Pro Glu
                230                 235                 240
Ala Asn Gly Ser Ala Gly Lys Arg Ser Lys Lys Lys Lys Gln Arg
                245                 250                 255
Lys Asp Ser Ala Ser Glu Glu Glu Ala Arg Val Gly Ala Gly Lys
                260                 265                 270
Arg Lys Arg Arg His Ser Glu Val Glu Thr Asp Ser Lys Lys Lys
                275                 280                 285
Lys Met Lys Leu Pro Glu His Pro Glu Gly Gly Glu Pro Glu Asp
                290                 295                 300
Asp Glu Ala Pro Ala Lys Gly Lys Phe Asn Trp Lys Gly Thr Ile
                305                 310                 315
Lys Ala Ile Leu Lys Gln Ala Pro Asp Asn Glu Ile Thr Ile Lys
                320                 325                 330
Lys Leu Arg Lys Lys Val Leu Ala Gln Tyr Tyr Thr Val Thr Asp
                335                 340                 345
Glu His His Arg Ser Glu Glu Glu Leu Leu Val Ile Phe Asn Lys
                350                 355                 360
Lys Ile Ser Lys Asn Pro Thr Phe Lys Leu Leu Lys Asp Lys Val
                365                 370                 375
Lys Leu Val Lys
```

<210> 2
<211> 136
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 815856

<400> 2

```
Met Phe Gly Thr Pro Gln Glu His Arg Asn Met Pro Gln Ala Asp
  1               5                  10                  15
Ala Met Val Leu Val Ala Arg Asn Tyr Glu Arg Tyr Lys Asn Glu
                20                  25                  30
Cys Arg Glu Lys Glu Arg Glu Glu Ile Ala Arg Gln Ala Ala Lys
                35                  40                  45
Met Ala Asp Glu Ala Ile Leu Gln Glu Arg Glu Arg Gly Gly Pro
                50                  55                  60
Glu Glu Gly Val Arg Gly Gly His Pro Pro Ala Ile Gln Ser Leu
```

```
                  65                  70                  75
Ile Asn Leu Leu Ala Asp Asn Arg Tyr Leu Thr Ala Glu Glu Thr
                  80                  85                  90
Asp Lys Ile Ile Asn Tyr Leu Arg Glu Arg Lys Glu Arg Leu Met
                  95                 100                 105
Arg Ser Ser Thr Asp Ser Leu Pro Gly Glu Leu Arg Gly Arg Ala
                 110                 115                 120
Glu Ala Arg Phe Pro Ala Asn His Ser Gly Arg Pro Arg Val Pro
                 125                 130                 135
Arg
```

<210> 3
<211> 230
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 996352

<400> 3

```
Met Ser Ser Ser Tyr Tyr Val Asn Ala Leu Phe Ser Lys Tyr Thr
  1               5                  10                  15
Ala Gly Ala Ser Leu Phe Gln Asn Ala Glu Pro Thr Ser Cys Ser
                 20                  25                  30
Phe Ala Pro Asn Ser Gln Arg Ser Gly Tyr Gly Ala Gly Ala Gly
                 35                  40                  45
Ala Phe Ala Ser Thr Val Pro Gly Leu Tyr Asn Val Asn Ser Pro
                 50                  55                  60
Leu Tyr Gln Ser Pro Phe Ala Ser Gly Tyr Gly Leu Gly Ala Asp
                 65                  70                  75
Ala Tyr Gly Asn Leu Pro Cys Ala Ser Tyr Asp Gln Asn Ile Pro
                 80                  85                  90
Gly Leu Cys Ser Asp Leu Ala Lys Gly Ala Cys Asp Lys Thr Asp
                 95                 100                 105
Glu Gly Ala Leu His Gly Ala Ala Glu Ala Asn Phe Arg Ile Tyr
                110                 115                 120
Pro Trp Met Arg Ser Ser Gly Pro Asp Arg Lys Arg Gly Arg Gln
                125                 130                 135
Thr Tyr Thr Arg Tyr Gln Thr Leu Glu Leu Glu Lys Glu Phe His
                140                 145                 150
Phe Asn Arg Tyr Leu Ile Arg Arg Arg Arg Ile Glu Ile Ala His
                155                 160                 165
Ala Leu Cys Leu Thr Glu Arg Gln Ile Lys Ile Trp Phe Gln Asn
                170                 175                 180
Arg Arg Met Lys Trp Lys Lys Glu His Lys Asp Glu Gly Pro Thr
                185                 190                 195
Ala Ala Ala Ala Pro Glu Gly Ala Val Pro Ser Ala Ala Ala Thr
                200                 205                 210
Ala Ala Ala Asp Lys Ala Asp Glu Glu Asp Asp Asp Glu Glu Glu
                215                 220                 225
Glu Asp Glu Glu Glu
                230
```

<210> 4
<211> 131
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 1273778

<400> 4
Met Ser Gln Val Thr Phe Ser Asp Val Ala Ile Asp Phe Ser His
1 5 10 15
Glu Glu Trp Ala Cys Leu Asp Ser Ala Gln Arg Asp Leu Tyr Lys
20 25 30
Asp Val Met Val Gln Asn Tyr Glu Asn Leu Val Ser Val Gly Leu
35 40 45
Ser Val Thr Lys Pro Tyr Val Ile Met Leu Leu Glu Asp Gly Lys
50 55 60
Glu Pro Trp Met Met Glu Lys Lys Leu Ser Lys Ala Tyr Pro Phe
65 70 75
Pro Leu Ser His Ser Val Pro Ala Ser Val Asn Phe Gly Phe Ser
80 85 90
Ala Leu Phe Glu His Cys Ser Glu Val Thr Glu Ile Phe Glu Leu
95 100 105
Ser Glu Leu Cys Val Phe Trp Val Leu His Phe Leu Ser Asn Ser
110 115 120
Pro Asn Ser Thr Val Glu Ala Phe Phe Lys Lys
125 130

<210> 5
<211> 411
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 1509715

<400> 5
Met Ser Lys Arg Pro Ser Tyr Ala Pro Pro Pro Thr Pro Ala Pro
1 5 10 15
Ala Thr Gln Met Pro Ser Thr Pro Gly Phe Val Gly Tyr Asn Pro
20 25 30
Tyr Ser His Leu Ala Tyr Asn Asn Tyr Arg Leu Gly Gly Asn Pro
35 40 45
Gly Thr Asn Ser Arg Val Thr Ala Ser Ser Gly Ile Thr Ile Pro
50 55 60
Lys Pro Pro Lys Pro Pro Asp Lys Pro Leu Met Pro Tyr Met Arg
65 70 75
Tyr Ser Arg Lys Val Trp Asp Gln Val Lys Ala Ser Asn Pro Asp
80 85 90
Leu Lys Leu Trp Glu Ile Gly Lys Ile Ile Gly Gly Met Trp Arg

```
                 95                  100                 105
Asp Leu Thr Asp Glu Glu Lys Gln Glu Tyr Leu Asn Glu Tyr Glu
                110                 115                 120
Ala Glu Lys Ile Glu Tyr Asn Glu Ser Met Lys Ala Tyr His Asn
                125                 130                 135
Ser Pro Ala Tyr Leu Ala Tyr Ile Asn Ala Lys Ser Arg Ala Glu
                140                 145                 150
Ala Ala Leu Glu Glu Glu Ser Arg Gln Arg Gln Ser Arg Met Glu
                155                 160                 165
Lys Gly Glu Pro Tyr Met Ser Ile Gln Pro Ala Glu Asp Pro Asp
                170                 175                 180
Asp Tyr Asp Asp Gly Phe Ser Met Lys His Thr Ala Thr Ala Arg
                185                 190                 195
Phe Gln Arg Asn His Arg Leu Ile Ser Glu Ile Leu Ser Glu Ser
                200                 205                 210
Val Val Pro Asp Val Arg Ser Val Val Thr Thr Ala Arg Met Gln
                215                 220                 225
Val Leu Lys Arg Gln Val Gln Ser Leu Met Val His Gln Arg Lys
                230                 235                 240
Leu Glu Ala Glu Leu Leu Gln Ile Glu Glu Arg His Gln Glu Lys
                245                 250                 255
Lys Arg Lys Phe Leu Glu Ser Thr Asp Ser Phe Asn Asn Glu Leu
                260                 265                 270
Lys Arg Leu Cys Gly Leu Lys Val Glu Val Asp Met Glu Lys Ile
                275                 280                 285
Ala Ala Glu Ile Ala Gln Ala Glu Glu Gln Ala Arg Lys Arg Gln
                290                 295                 300
Glu Glu Arg Glu Lys Glu Ala Ala Glu Gln Ala Glu Arg Ser Gln
                305                 310                 315
Ser Ser Ile Val Pro Glu Glu Glu Gln Ala Ala Asn Lys Gly Glu
                320                 325                 330
Glu Lys Lys Asp Asp Glu Asn Ile Pro Met Glu Thr Glu Glu Thr
                335                 340                 345
His Leu Glu Glu Thr Thr Glu Ser Gln Gln Asn Gly Glu Glu Gly
                350                 355                 360
Thr Ser Thr Pro Glu Asp Lys Glu Ser Gly Gln Glu Gly Val Asp
                365                 370                 375
Ser Met Ala Glu Glu Gly Thr Ser Asp Ser Asn Thr Gly Ser Glu
                380                 385                 390
Ser Asn Ser Ala Thr Val Glu Glu Pro Pro Thr Asp Pro Ile Pro
                395                 400                 405
Glu Asp Glu Lys Lys Glu
                410
```

<210> 6
<211> 226
<212> PRT
<213> Homo sapiens
<220>
<221> unsure
<222> 221, 228
<223> unknown or other

<220>

<221> misc_feature
<223> Incyte Clone No.: 1676367

<400> 6

```
Met Ala Ala Lys Val Asp Leu Ser Thr Ser Thr Asp Trp Lys Glu
  1               5                  10                  15
Ala Lys Ser Phe Leu Lys Gly Leu Ser Asp Lys Gln Arg Glu Glu
                 20                  25                  30
His Tyr Phe Cys Lys Asp Phe Val Arg Leu Lys Lys Ile Pro Thr
                 35                  40                  45
Trp Lys Glu Met Ala Lys Gly Val Ala Val Lys Val Glu Glu Pro
                 50                  55                  60
Arg Tyr Lys Lys Asp Lys Gln Leu Asn Glu Lys Ile Ser Leu Leu
                 65                  70                  75
Arg Ser Asp Ile Thr Lys Leu Glu Val Asp Ala Ile Val Asn Ala
                 80                  85                  90
Ala Asn Ser Ser Leu Leu Gly Gly Gly Gly Val Asp Gly Cys Ile
                 95                 100                 105
His Arg Ala Ala Gly Pro Leu Leu Thr Asp Glu Cys Arg Thr Leu
                110                 115                 120
Gln Ser Cys Lys Thr Gly Lys Ala Lys Ile Thr Gly Gly Tyr Arg
                125                 130                 135
Leu Pro Ala Lys Tyr Val Ile His Thr Val Gly Pro Ile Ala Tyr
                140                 145                 150
Gly Glu Pro Ser Ala Ser Gln Ala Ala Glu Leu Arg Ser Cys Tyr
                155                 160                 165
Leu Ser Ser Leu Asp Leu Leu Leu Glu His Arg Leu Arg Ser Val
                170                 175                 180
Ala Phe Pro Cys Ile Ser Thr Gly Val Phe Gly Tyr Pro Cys Glu
                185                 190                 195
Ala Ala Ala Glu Ile Val Leu Ala Thr Leu Arg Glu Trp Leu Gly
                200                 205                 210
Ser Ser Thr Arg Glu Pro Arg Xaa Asn Leu Asn Phe Xaa Glu Pro
                215                 220                 225
Gly
```

<210> 7
<211> 183
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 1734119

<400> 7

```
Met Gly Arg Leu Cys Cys Leu Arg Pro Pro Pro His Arg Asp Pro
  1               5                  10                  15
Ala Arg Leu Leu Leu Ala Ser Thr Asp Asp Lys Arg Asn Ser Pro
                 20                  25                  30
Lys Ile Arg Pro Leu Gln Pro Ala Val Pro Ala Cys Leu Pro Ala
                 35                  40                  45
Thr Val Arg Pro Ala Leu Ala Ser Ser Ser Ala Gly Leu Ser Ala
                 50                  55                  60
```

```
Gly Phe Trp Gly Gln Lys Ser Gly Glu Pro Arg Gly Arg Val Arg
                65                  70                  75
Gly Asp Gln Val Arg Ala Ala Thr Phe Leu Val Ile Ser Pro Met
                80                  85                  90
Gly Arg Arg Gly Trp Arg Asp Thr Ala Pro Pro Gly Phe Pro Thr
                95                 100                 105
Pro Leu Leu Ser His Pro Glu Ala Ser Phe Phe Cys Ala Arg Cys
               110                 115                 120
Leu Pro Lys Arg Val Gly Ala Arg Ser Pro Pro Trp Arg Val Leu
               125                 130                 135
Gly Pro Gly Gly Ala Leu Gly Glu Gln Met Gly Pro Pro Leu Ala
               140                 145                 150
Gly Pro Leu Gln Leu Phe Pro Ala Ala Glu Pro Ser Gly Gly Pro
               155                 160                 165
Val Leu Val Ala Ser Leu Arg Ala Gln Ile Ala Gln Gly Asp Leu
               170                 175                 180
Ala Val Ala
```

<210> 8
<211> 317
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 1944813

<400> 8

```
Met Lys Ser Asp Cys Met Gln Thr Thr Ile Cys Gln Glu Arg Lys
  1                 5                  10                  15
Lys Asp Pro Ile Glu Met Phe His Ser Gly Gln Leu Val Lys Val
                20                  25                  30
Cys Ala Pro Met Val Arg Tyr Ser Lys Leu Ala Phe Arg Thr Leu
                35                  40                  45
Val Arg Lys Tyr Ser Cys Asp Leu Cys Tyr Thr Pro Met Ile Val
                50                  55                  60
Ala Ala Asp Phe Val Lys Ser Ile Lys Ala Arg Asp Ser Glu Phe
                65                  70                  75
Thr Thr Asn Gln Gly Asp Cys Pro Leu Ile Val Gln Phe Ala Ala
                80                  85                  90
Asn Asp Ala Arg Leu Leu Ser Asp Ala Ala Arg Ile Val Cys Pro
                95                 100                 105
Tyr Ala Asn Gly Ile Asp Ile Asn Cys Gly Cys Pro Gln Arg Trp
               110                 115                 120
Ala Met Ala Glu Gly Tyr Gly Ala Cys Leu Ile Asn Lys Pro Glu
               125                 130                 135
Leu Val Gln Asp Met Val Lys Gln Val Arg Asn Gln Val Glu Thr
               140                 145                 150
Pro Gly Phe Ser Val Ser Ile Lys Ile Arg Ile His Asp Asp Leu
               155                 160                 165
Lys Arg Thr Val Asp Leu Cys Gln Lys Ala Glu Ala Thr Gly Val
               170                 175                 180
Ser Trp Ile Thr Val His Gly Arg Thr Ala Glu Glu Arg His Gln
               185                 190                 195
```

```
Pro Val His Tyr Asp Ser Ile Lys Ile Ile Lys Glu Asn Met Ser
                200                 205                 210
Ile Pro Val Ile Ala Asn Gly Asp Ile Arg Ser Leu Lys Glu Ala
                215                 220                 225
Glu Asn Val Trp Arg Ile Thr Gly Thr Asp Gly Val Met Val Ala
                230                 235                 240
Arg Gly Leu Leu Ala Asn Pro Ala Met Phe Ala Gly Tyr Glu Glu
                245                 250                 255
Thr Pro Leu Lys Cys Ile Trp Asp Trp Val Asp Ile Ala Leu Glu
                260                 265                 270
Leu Gly Thr Pro Tyr Met Cys Phe His Gln His Leu Met Tyr Met
                275                 280                 285
Met Glu Lys Ile Thr Ser Arg Gln Glu Lys Arg Val Phe Asn Ala
                290                 295                 300
Leu Ser Ser Thr Ser Ala Ile Ile Asp Tyr Leu Thr Asp His Tyr
                305                 310                 315
Gly Ile
```

<210> 9
<211> 479
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 2683322

<400> 9

```
Met Ala Thr Asp Ser Trp Ala Leu Ala Val Asp Glu Gln Glu Ala
  1                 5                  10                  15
Ala Ala Glu Ser Leu Ser Asn Leu His Leu Lys Glu Glu Lys Ile
                 20                  25                  30
Lys Pro Asp Thr Asn Gly Ala Val Val Lys Thr Asn Ala Asn Ala
                 35                  40                  45
Glu Lys Thr Asp Glu Glu Glu Lys Glu Asp Arg Ala Ala Gln Ser
                 50                  55                  60
Leu Leu Asn Lys Leu Ile Arg Ser Asn Leu Val Asp Asn Thr Asn
                 65                  70                  75
Gln Val Glu Val Leu Gln Arg Asp Pro Asn Ser Pro Leu Tyr Ser
                 80                  85                  90
Val Lys Ser Phe Glu Glu Leu Arg Leu Lys Pro Gln Leu Leu Gln
                 95                 100                 105
Gly Val Tyr Ala Met Gly Phe Asn Arg Pro Ser Lys Ile Gln Glu
                110                 115                 120
Asn Ala Leu Pro Met Met Leu Ala Glu Pro Pro Gln Asn Leu Ile
                125                 130                 135
Ala Gln Ser Gln Ser Gly Thr Gly Lys Thr Ala Ala Phe Val Leu
                140                 145                 150
Ala Met Leu Ser Arg Val Glu Pro Ser Asp Arg Tyr Pro Gln Cys
                155                 160                 165
Leu Cys Leu Ser Pro Thr Tyr Glu Leu Ala Leu Gln Thr Gly Lys
                170                 175                 180
Val Ile Glu Gln Met Gly Lys Phe Tyr Pro Glu Leu Lys Leu Ala
                185                 190                 195
```

```
Tyr Ala Val Arg Gly Asn Lys Leu Glu Arg Gly Gln Lys Ile Ser
                200                 205                 210
Glu Gln Ile Val Ile Gly Thr Pro Gly Thr Val Leu Asp Trp Cys
                215                 220                 225
Ser Lys Leu Lys Phe Ile Asp Pro Lys Lys Ile Lys Val Phe Val
                230                 235                 240
Leu Asp Glu Ala Asp Val Met Ile Ala Thr Gln Gly His Gln Asp
                245                 250                 255
Gln Ser Ile Arg Ile Gln Arg Met Leu Pro Arg Asn Cys Gln Met
                260                 265                 270
Leu Leu Phe Ser Ala Thr Phe Glu Asp Ser Val Trp Lys Phe Ala
                275                 280                 285
Gln Lys Val Val Pro Asp Pro Asn Val Ile Lys Leu Lys Arg Glu
                290                 295                 300
Glu Glu Thr Leu Asp Thr Ile Lys Gln Tyr Tyr Val Leu Cys Ser
                305                 310                 315
Ser Arg Asp Glu Lys Phe Gln Ala Leu Cys Asn Leu Tyr Gly Ala
                320                 325                 330
Ile Thr Ile Ala Gln Ala Met Ile Phe Cys His Thr Arg Lys Thr
                335                 340                 345
Ala Ser Trp Leu Ala Ala Glu Leu Ser Lys Glu Gly His Gln Val
                350                 355                 360
Ala Leu Leu Ser Gly Glu Met Met Val Glu Gln Arg Ala Ala Val
                365                 370                 375
Ile Glu Arg Phe Arg Glu Gly Lys Glu Lys Val Leu Val Thr Thr
                380                 385                 390
Asn Val Cys Ala Arg Gly Ile Asp Val Glu Gln Val Ser Val Val
                395                 400                 405
Ile Asn Phe Asp Leu Pro Val Asp Lys Asp Gly Asn Pro Asp Asn
                410                 415                 420
Glu Thr Tyr Leu His Arg Ile Gly Arg Thr Gly Arg Phe Gly Lys
                425                 430                 435
Arg Gly Leu Ala Val Asn Met Val Asp Ser Lys His Ser Met Asn
                440                 445                 450
Ile Leu Asn Arg Ile Gln Glu His Phe Asn Lys Lys Ile Glu Arg
                455                 460                 465
Leu Asp Thr Asp Asp Leu Asp Glu Ile Glu Lys Ile Ala Asn
                470                 475
```

<210> 10
<211> 582
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 2684552

```
<400> 10
Met Ala Glu Phe Leu Asp Asp Gln Glu Thr Arg Leu Cys Asp Asn
  1               5                  10                  15
Cys Lys Lys Glu Ile Pro Val Phe Asn Phe Thr Ile His Glu Ile
                20                  25                  30
His Cys Gln Arg Asn Ile Gly Met Cys Pro Thr Cys Lys Glu Pro
```

```
                35                  40                  45
Phe Pro Lys Ser Asp Met Glu Thr His Met Ala Ala Glu His Cys
                50                  55                  60
Gln Val Thr Cys Lys Cys Asn Lys Lys Leu Glu Lys Arg Leu Leu
                65                  70                  75
Lys Lys His Glu Glu Thr Glu Cys Pro Leu Arg Leu Ala Val Cys
                80                  85                  90
Gln His Cys Asp Leu Glu Leu Ser Ile Leu Lys Leu Lys Glu His
                95                 100                 105
Glu Asp Tyr Cys Gly Ala Arg Thr Glu Leu Cys Gly Asn Cys Gly
               110                 115                 120
Arg Asn Val Leu Val Lys Asp Leu Lys Thr His Pro Glu Val Cys
               125                 130                 135
Gly Arg Glu Gly Glu Glu Lys Arg Asn Glu Val Ala Ile Pro Pro
               140                 145                 150
Asn Ala Tyr Asp Glu Ser Trp Gly Gln Asp Gly Ile Trp Ile Ala
               155                 160                 165
Ser Gln Leu Leu Arg Gln Ile Glu Ala Leu Asp Pro Pro Met Arg
               170                 175                 180
Leu Pro Arg Arg Pro Leu Arg Ala Phe Glu Ser Asp Val Phe His
               185                 190                 195
Asn Arg Thr Thr Asn Gln Arg Asn Ile Thr Ala Gln Val Ser Ile
               200                 205                 210
Gln Asn Asn Leu Phe Glu Glu Gln Glu Arg Gln Glu Arg Asn Arg
               215                 220                 225
Gly Gln Gln Pro Pro Lys Glu Gly Gly Glu Glu Ser Ala Asn Leu
               230                 235                 240
Asp Phe Met Leu Ala Leu Ser Leu Gln Asn Glu Gly Gln Ala Ser
               245                 250                 255
Ser Val Ala Glu Gln Asp Phe Trp Arg Ala Val Cys Glu Ala Asp
               260                 265                 270
Gln Ser His Gly Gly Pro Arg Ser Leu Ser Asp Ile Lys Gly Ala
               275                 280                 285
Ala Asp Glu Ile Met Leu Pro Cys Glu Phe Cys Glu Glu Leu Tyr
               290                 295                 300
Pro Glu Glu Leu Leu Ile Asp His Gln Thr Ser Cys Asn Pro Ser
               305                 310                 315
Arg Ala Leu Pro Ser Leu Asn Thr Gly Ser Ser Ser Pro Arg Gly
               320                 325                 330
Val Glu Glu Pro Asp Val Ile Phe Gln Asn Phe Leu Gln Gln Ala
               335                 340                 345
Ala Ser Asn Gln Leu Asp Ser Leu Met Gly Leu Ser Asn Ser His
               350                 355                 360
Pro Val Glu Glu Ser Ile Ile Ile Pro Cys Glu Phe Cys Gly Val
               365                 370                 375
Gln Leu Glu Glu Glu Val Leu Phe His His Gln Asp Gln Cys Asp
               380                 385                 390
Gln Arg Pro Ala Thr Ala Thr Asn His Val Thr Glu Gly Ile Pro
               395                 400                 405
Arg Leu Asp Ser Gln Pro Gln Glu Thr Ser Pro Glu Leu Pro Arg
               410                 415                 420
Arg Arg Val Arg His Gln Gly Asp Leu Ser Ser Gly Tyr Leu Asp
               425                 430                 435
Asp Thr Lys Gln Glu Thr Ala Asn Gly Pro Thr Ser Cys Leu Pro
               440                 445                 450
Pro Ser Arg Pro Ile Asn Asn Met Thr Ala Thr Tyr Asn Gln Leu
               455                 460                 465
```

```
Ser Arg Ser Thr Ser Gly Pro Arg Pro Gly Cys Gln Pro Ser Ser
                470                 475                 480
Pro Cys Val Pro Lys Leu Ser Asn Ser Asp Ser Gln Asp Ile Gln
                485                 490                 495
Gly Arg Asn Arg Asp Ser Gln Asn Gly Ala Ile Ala Pro Gly His
                500                 505                 510
Val Ser Val Ile Arg Pro Pro Gln Asn Leu Tyr Pro Glu Asn Ile
                515                 520                 525
Val Pro Ser Phe Ser Pro Gly Pro Ser Gly Arg Tyr Gly Ala Ser
                530                 535                 540
Gly Arg Ser Glu Gly Gly Arg Asn Ser Arg Val Thr Pro Ala Ala
                545                 550                 555
Ala Asn Tyr Arg Ser Arg Thr Ala Lys Ala Lys Pro Ser Lys Gln
                560                 565                 570
Gln Gly Ala Gly Asp Ala Glu Glu Glu Glu Glu Glu
                575                 580
```

<210> 11
<211> 327
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 2830310

<400> 11

```
Met Arg Trp Pro Gly His Tyr Ser Arg Ala Pro Tyr Pro Tyr Phe
  1               5                  10                  15
Ser Ser Arg His Phe Ser Leu Asn Trp Arg Pro Pro Cys Leu Phe
                20                  25                  30
Glu Ser Arg Thr Gln Phe Gln Tyr Cys Asn Trp Arg Pro Asp Asn
                35                  40                  45
Leu Ser Gln Thr Ser Leu Ile His Leu Ser Ser Tyr Val Met Asn
                50                  55                  60
Ala Glu Gly Asp Glu Pro Ser Ser Lys Arg Arg Lys His Gln Gly
                65                  70                  75
Val Ile Lys Arg Asn Trp Glu Tyr Ile Cys Ser His Asp Lys Glu
                80                  85                  90
Lys Thr Lys Ile Leu Gly Asp Lys Asn Val Asp Pro Lys Cys Glu
                95                 100                 105
Asp Ser Glu Asn Lys Phe Asp Phe Ser Val Met Ser Tyr Asn Ile
               110                 115                 120
Leu Ser Gln Asp Leu Leu Glu Asp Asn Ser His Leu Tyr Arg His
               125                 130                 135
Cys Arg Arg Pro Val Leu His Trp Ser Phe Arg Phe Pro Asn Ile
               140                 145                 150
Leu Lys Glu Ile Lys His Phe Asp Ala Asp Val Leu Cys Leu Gln
               155                 160                 165
Glu Val Gln Glu Asp His Tyr Gly Ala Glu Ile Arg Pro Ser Leu
               170                 175                 180
Glu Ser Leu Gly Tyr His Cys Glu Tyr Lys Met Arg Thr Gly Arg
               185                 190                 195
Lys Pro Asp Gly Cys Ala Ile Cys Phe Lys His Ser Lys Phe Ser
```

```
                200                 205                 210
Leu Leu Ser Val Asn Pro Val Glu Phe Phe Arg Pro Asp Ile Ser
                215                 220                 225
Leu Leu Asp Arg Asp Asn Val Gly Leu Val Leu Leu Leu Gln Pro
                230                 235                 240
Lys Ile Pro Tyr Ala Ala Cys Pro Ala Ile Cys Val Ala Asn Thr
                245                 250                 255
His Leu Leu Tyr Asn Pro Arg Arg Gly Asp Ile Lys Leu Thr Gln
                260                 265                 270
Leu Ala Met Leu Leu Ala Glu Ile Ser Ser Val Ala His Gln Lys
                275                 280                 285
Asp Gly Ser Phe Cys Pro Ile Val Met Cys Gly Asp Phe Asn Ser
                290                 295                 300
Val Pro Gly Ser Pro Leu Tyr Ser Phe Ile Lys Glu Gly Lys Leu
                305                 310                 315
Asn Tyr Glu Gly Leu Pro Ile Gly Lys Thr Val Ile
                320                 325
```

<210> 12
<211> 502
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 2963346

<400> 12

```
Met Ala Ser Lys Lys Leu Gly Ala Asp Phe His Gly Thr Phe Ser
  1                 5                  10                  15
Tyr Leu Asp Asp Val Pro Phe Lys Thr Gly Asp Lys Phe Lys Thr
                 20                  25                  30
Pro Ala Lys Val Gly Leu Pro Ile Gly Phe Ser Leu Pro Asp Cys
                 35                  40                  45
Leu Gln Val Val Arg Glu Val Gln Tyr Asp Phe Ser Leu Glu Lys
                 50                  55                  60
Lys Thr Ile Glu Trp Ala Glu Glu Ile Lys Lys Ile Glu Glu Ala
                 65                  70                  75
Glu Arg Glu Ala Glu Cys Lys Ile Ala Glu Ala Glu Ala Lys Val
                 80                  85                  90
Asn Ser Lys Ser Gly Pro Glu Gly Asp Ser Lys Met Ser Phe Ser
                 95                 100                 105
Lys Thr His Ser Thr Ala Thr Met Pro Pro Pro Ile Asn Pro Ile
                110                 115                 120
Leu Ala Ser Leu Gln His Asn Ser Ile Leu Thr Pro Thr Arg Val
                125                 130                 135
Ser Ser Ser Ala Thr Lys Gln Lys Val Leu Ser Pro Pro His Ile
                140                 145                 150
Lys Ala Asp Phe Asn Leu Ala Asp Phe Glu Cys Glu Glu Asp Pro
                155                 160                 165
Phe Asp Asn Leu Glu Leu Lys Thr Ile Asp Glu Lys Glu Glu Leu
                170                 175                 180
Arg Asn Ile Leu Val Gly Thr Thr Gly Pro Ile Met Ala Gln Leu
                185                 190                 195
```

```
Leu Asp Asn Asn Leu Pro Arg Gly Gly Ser Gly Ser Val Leu Gln
                200                 205                 210
Asp Glu Glu Val Leu Ala Ser Leu Glu Arg Ala Thr Leu Asp Phe
                215                 220                 225
Lys Pro Leu His Lys Pro Asn Gly Phe Ile Thr Leu Pro Gln Leu
                230                 235                 240
Gly Asn Cys Glu Lys Met Ser Leu Ser Ser Lys Val Ser Leu Pro
                245                 250                 255
Pro Ile Pro Ala Val Ser Asn Ile Lys Ser Leu Ser Phe Pro Lys
                260                 265                 270
Leu Asp Ser Asp Asp Ser Asn Gln Lys Thr Ala Lys Leu Ala Ser
                275                 280                 285
Thr Phe His Ser Thr Ser Cys Leu Arg Asn Gly Thr Phe Gln Asn
                290                 295                 300
Ser Leu Lys Pro Ser Thr Gln Ser Ser Ala Ser Glu Leu Asn Gly
                305                 310                 315
His His Thr Leu Gly Leu Ser Ala Leu Asn Leu Asp Ser Gly Thr
                320                 325                 330
Glu Met Pro Ala Leu Thr Ser Ser Gln Met Pro Ser Leu Ser Val
                335                 340                 345
Leu Ser Val Cys Thr Glu Glu Ser Ser Pro Pro Asn Thr Gly Pro
                350                 355                 360
Thr Val Thr Pro Pro Asn Phe Ser Val Ser Gln Val Pro Asn Met
                365                 370                 375
Pro Ser Cys Pro Gln Ala Tyr Ser Glu Leu Gln Met Leu Ser Pro
                380                 385                 390
Ser Glu Arg Gln Cys Val Glu Thr Val Val Asn Met Gly Tyr Ser
                395                 400                 405
Tyr Glu Cys Val Leu Arg Ala Met Lys Lys Lys Gly Glu Asn Ile
                410                 415                 420
Glu Gln Ile Leu Asp Tyr Leu Phe Ala His Gly Gln Leu Cys Glu
                425                 430                 435
Lys Gly Phe Asp Pro Leu Leu Val Glu Glu Ala Leu Glu Met His
                440                 445                 450
Gln Cys Ser Glu Glu Lys Met Met Glu Phe Leu Gln Leu Met Ser
                455                 460                 465
Lys Phe Lys Glu Met Gly Phe Glu Leu Lys Asp Ile Lys Glu Val
                470                 475                 480
Leu Leu Leu His Asn Asn Asp Gln Asp Asn Ala Leu Glu Asp Leu
                485                 490                 495
Met Ala Arg Ala Gly Ala Ser
                500
```

<210> 13
<211> 375
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 2994234

<400> 13
Met Leu Trp Lys Leu Leu Leu Arg Ser Gln Ser Cys Arg Leu Cys

```
  1               5                  10                  15
Ser Phe Arg Lys Met Arg Ser Pro Pro Lys Tyr Arg Pro Phe Leu
            20                  25                  30
Ala Cys Phe Thr Tyr Thr Thr Asp Lys Gln Ser Ser Lys Glu Asn
            35                  40                  45
Thr Arg Thr Val Glu Lys Leu Tyr Lys Cys Ser Val Asp Ile Arg
            50                  55                  60
Lys Ile Arg Arg Leu Lys Gly Trp Val Leu Leu Glu Asp Glu Thr
            65                  70                  75
Tyr Val Glu Glu Ile Ala Asn Ile Leu Gln Glu Leu Gly Ala Asp
            80                  85                  90
Glu Thr Ala Val Ala Ser Ile Leu Glu Arg Cys Pro Glu Ala Ile
            95                  100                 105
Val Cys Ser Pro Thr Ala Val Asn Thr Gln Arg Lys Leu Trp Gln
            110                 115                 120
Leu Val Cys Lys Asn Glu Glu Glu Leu Ile Lys Leu Ile Glu Gln
            125                 130                 135
Phe Pro Glu Ser Phe Phe Thr Ile Lys Asp Gln Glu Asn Gln Lys
            140                 145                 150
Leu Asn Val Gln Phe Phe Gln Glu Leu Gly Leu Lys Asn Val Val
            155                 160                 165
Ile Ser Arg Leu Leu Thr Ala Ala Pro Asn Val Phe His Asn Pro
            170                 175                 180
Val Glu Lys Asn Lys Gln Met Val Arg Ile Leu Gln Glu Ser Tyr
            185                 190                 195
Leu Asp Val Gly Gly Ser Glu Ala Asn Met Lys Val Trp Leu Leu
            200                 205                 210
Lys Leu Leu Ser Gln Asn Pro Phe Ile Leu Leu Asn Ser Pro Thr
            215                 220                 225
Ala Ile Lys Glu Thr Leu Glu Phe Leu Gln Glu Gln Gly Phe Thr
            230                 235                 240
Ser Phe Glu Ile Leu Gln Leu Leu Ser Lys Leu Lys Gly Phe Leu
            245                 250                 255
Phe Gln Leu Cys Pro Arg Ser Ile Gln Asn Ser Ile Ser Phe Ser
            260                 265                 270
Lys Asn Ala Phe Lys Cys Thr Asp His Asp Leu Lys Gln Leu Val
            275                 280                 285
Leu Lys Cys Pro Ala Leu Leu Tyr Tyr Ser Val Pro Val Leu Glu
            290                 295                 300
Glu Arg Met Gln Gly Leu Leu Arg Glu Gly Ile Ser Ile Ala Gln
            305                 310                 315
Ile Arg Glu Thr Pro Met Val Leu Glu Leu Thr Pro Gln Ile Val
            320                 325                 330
Gln Tyr Arg Ile Arg Lys Leu Asn Ser Ser Gly Tyr Arg Ile Lys
            335                 340                 345
Asp Gly His Leu Ala Asn Leu Asn Gly Ser Lys Lys Glu Phe Glu
            350                 355                 360
Ala Asn Phe Gly Lys Ile Gln Ala Lys Lys Ser Lys Ala Ile Ile
            365                 370                 375
```

<210> 14
<211> 341
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 4115958

<400> 14

```
Met His Asp Ser Ser Ser Val Ala Ser Lys Val Phe Arg Ser Ser
  1               5                  10                  15
Tyr Glu Asp Lys Asn Leu Leu Lys Lys Asn Lys Asp Glu Ser Ser
             20                  25                  30
Val Ser Ile Ser His Thr Lys Cys Ser Leu Leu Gly Asp Ile Ser
             35                  40                  45
Asp Gly Lys Asn Leu Ile Pro Asn Lys Cys Phe Thr Ser Phe Lys
             50                  55                  60
Asn Asn Ser Lys Glu Lys Cys Ser Leu Lys His Gln Thr Arg Asn
             65                  70                  75
Gln Cys Gln Asn Asn Pro Ser Glu Ile Ile Gln Ser Thr Tyr Gln
             80                  85                  90
Glu Thr Gln Asn Lys Ser Ser Ser Leu Ser Thr Ser Ser Ile Leu
             95                 100                 105
Ser Gln His Lys Glu Asn Asn Leu Asp Leu Thr Ser Arg Phe Lys
            110                 115                 120
Glu Gln Glu Met Ser Asn Gly Ile Asp Lys Gln Tyr Ser Asn Cys
            125                 130                 135
Thr Thr Ile Asp Lys Gln Ile Cys Thr Asn Lys Tyr Lys Glu Lys
            140                 145                 150
Ile Ile Asn Glu Asn Tyr Asn Pro Lys Phe Phe Gly Asn Leu Gln
            155                 160                 165
Ser Asp Asp Ser Lys Lys Asn Asp Ser Lys Ile Lys Val Thr Val
            170                 175                 180
Leu Glu Met Ser Glu Tyr Leu Asn Lys Tyr Glu Ser Met Ser Ser
            185                 190                 195
Asn Lys Asp Ser Lys Arg Pro Lys Thr Cys Glu Gln Asn Thr Gln
            200                 205                 210
Leu Asn Ser Ile Glu Asn Tyr Leu Asn Lys Asp Asn Glu Gly Phe
            215                 220                 225
Lys Cys Lys Lys Ser Asp Gln Leu Lys Asn Glu Gln Asp Lys Gln
            230                 235                 240
Glu Asp Pro Thr Asn Glu Lys Ser Gln Asn Tyr Ser Gln Arg Arg
            245                 250                 255
Ser Ile Lys Asp Cys Leu Ser Thr Cys Glu Gln Pro Lys Asn Thr
            260                 265                 270
Glu Val Leu Arg Thr Thr Leu Lys His Ser Asn Val Trp Arg Lys
            275                 280                 285
His Asn Phe His Ser Leu Asp Gly Thr Ser Thr Arg Ala Phe His
            290                 295                 300
Pro Gln Thr Gly Leu Pro Leu Leu Ser Ser Pro Val Pro Gln Arg
            305                 310                 315
Lys Thr Gln Ser Gly Cys Phe Asp Leu Asp Ser Ser Leu Leu His
            320                 325                 330
Leu Lys Ser Phe Ser Ser Arg Arg Asn Leu Ser
            335                 340
```

<210> 15
<211> 269

<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 779255

<400> 15

```
Met His Thr Glu Thr Ile Lys Pro His Lys Cys Pro His Cys Ser
  1               5                  10                  15
Lys Thr Phe Ala Asn Thr Ser Tyr Leu Ala Gln His Leu Arg Ile
             20                  25                  30
His Ser Gly Ala Lys Pro Tyr Asn Cys Ser Tyr Cys Gln Lys Ala
             35                  40                  45
Phe Arg Gln Leu Ser His Leu Gln Gln His Thr Arg Ile His Thr
             50                  55                  60
Gly Asp Arg Pro Tyr Lys Cys Ala His Pro Gly Cys Glu Lys Ala
             65                  70                  75
Phe Thr Gln Leu Ser Asn Leu Gln Ser His Arg Arg Gln His Asn
             80                  85                  90
Lys Asp Lys Pro Phe Lys Cys His Asn Cys His Arg Ala Tyr Thr
             95                 100                 105
Asp Ala Ala Ser Leu Glu Val His Leu Ser Thr His Thr Val Lys
            110                 115                 120
His Ala Lys Val Tyr Thr Cys Thr Ile Cys Ser Arg Ala Tyr Thr
            125                 130                 135
Ser Glu Thr Tyr Leu Met Lys His Met Arg Lys His Asn Pro Pro
            140                 145                 150
Asp Leu Gln Gln Gln Val Gln Ala Ala Ala Ala Ala Ala Ala Val
            155                 160                 165
Ala Gln Ala Gln Ala Gln Ala Gln Ala Gln Ala Gln Ala Gln Ala
            170                 175                 180
Gln Ala Gln Ala Gln Ala Gln Ala Ser Gln Ala Ser Gln Gln Gln
            185                 190                 195
Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Pro Pro
            200                 205                 210
Pro His Phe Gln Ser Pro Gly Ala Ala Pro Gln Gly Gly Gly Gly
            215                 220                 225
Gly Asp Ser Asn Pro Asn Pro Pro Pro Gln Cys Ser Phe Asp Leu
            230                 235                 240
Thr Pro Tyr Lys Thr Ala Glu His His Lys Asp Ile Cys Leu Thr
            245                 250                 255
Val Thr Thr Ser Thr Ile Gln Val Glu His Leu Ala Ser Ser
            260                 265
```

<210> 16
<211> 264
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 1303605

<400> 16

```
Met Glu Asn Tyr Gly Ile Glu Trp His Ser Val Arg Asp Ser Glu
  1               5                  10                  15
Gly Gln Lys Leu Leu Ile Gly Val Gly Pro Glu Gly Ile Ser Ile
                 20                  25                  30
Cys Lys Asp Asp Phe Ser Pro Ile Asn Arg Ile Ala Tyr Pro Val
                 35                  40                  45
Val Gln Met Ala Thr Gln Ser Gly Lys Asn Val Tyr Leu Thr Val
                 50                  55                  60
Thr Lys Glu Ser Gly Asn Ser Ile Val Leu Leu Phe Lys Met Ile
                 65                  70                  75
Ser Thr Arg Ala Ala Ser Gly Leu Tyr Arg Ala Ile Thr Glu Thr
                 80                  85                  90
His Ala Phe Tyr Arg Cys Asp Thr Val Thr Ser Ala Val Met Met
                 95                 100                 105
Gln Tyr Ser Arg Asp Leu Lys Gly His Leu Ala Ser Leu Phe Leu
                110                 115                 120
Asn Glu Asn Ile Asn Leu Gly Lys Lys Tyr Val Phe Asp Ile Lys
                125                 130                 135
Arg Thr Ser Lys Glu Val Tyr Asp His Ala Arg Arg Ala Leu Tyr
                140                 145                 150
Asn Ala Gly Val Val Asp Leu Val Ser Arg Ser Asn Gln Ser Pro
                155                 160                 165
Ser His Ser Pro Leu Lys Ser Ser Glu Ser Ser Met Asn Cys Ser
                170                 175                 180
Ser Cys Glu Gly Leu Ser Cys Gln Gln Thr Arg Val Leu Gln Glu
                185                 190                 195
Lys Leu Arg Lys Leu Lys Glu Ala Met Leu Cys Met Val Cys Cys
                200                 205                 210
Glu Glu Glu Ile Asn Ser Thr Phe Cys Pro Cys Gly His Thr Val
                215                 220                 225
Cys Cys Glu Ser Cys Ala Ala Gln Leu Gln Ser Cys Pro Val Cys
                230                 235                 240
Arg Ser Arg Val Glu His Val Gln His Val Tyr Leu Pro Thr His
                245                 250                 255
Thr Ser Leu Leu Asn Leu Thr Val Ile
                260
```

<210> 17
<211> 605
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 1611167

<400> 17

```
Met Ala Ala Gly Gln Arg Glu Ala Arg Pro Gln Val Ser Leu Thr
  1               5                  10                  15
Phe Glu Asp Val Ala Val Leu Phe Thr Arg Asp Glu Trp Arg Lys
                 20                  25                  30
Leu Ala Pro Ser Gln Arg Asn Leu Tyr Arg Asp Val Met Leu Glu
                 35                  40                  45
```

```
Asn Tyr Arg Asn Leu Val Ser Leu Gly Leu Pro Phe Thr Lys Pro
                50                  55                  60
Lys Val Ile Ser Leu Leu Gln Gln Gly Glu Asp Pro Trp Glu Val
                65                  70                  75
Glu Lys Asp Gly Ser Gly Val Ser Ser Leu Gly Ser Lys Ser Ser
                80                  85                  90
His Lys Thr Thr Lys Ser Thr Gln Thr Gln Asp Ser Ser Phe Gln
                95                 100                 105
Gly Leu Ile Leu Lys Arg Ser Asn Arg Asn Val Pro Trp Asp Leu
               110                 115                 120
Lys Leu Glu Lys Pro Tyr Ile Tyr Glu Gly Arg Leu Glu Lys Lys
               125                 130                 135
Gln Asp Lys Lys Gly Ser Phe Gln Ile Val Ser Ala Thr His Lys
               140                 145                 150
Lys Ile Pro Thr Ile Glu Arg Ser His Lys Asn Thr Glu Leu Ser
               155                 160                 165
Gln Asn Phe Ser Pro Lys Ser Val Leu Ile Arg Gln Gln Ile Leu
               170                 175                 180
Pro Arg Glu Lys Thr Pro Pro Lys Cys Glu Ile Gln Gly Asn Ser
               185                 190                 195
Leu Lys Gln Asn Ser Gln Leu Leu Asn Gln Pro Lys Ile Thr Ala
               200                 205                 210
Asp Lys Arg Tyr Lys Cys Ser Leu Cys Glu Lys Thr Phe Ile Asn
               215                 220                 225
Thr Ser Ser Leu Arg Lys His Glu Lys Asn His Ser Gly Glu Lys
               230                 235                 240
Leu Phe Lys Cys Lys Glu Cys Ser Lys Ala Phe Ser Gln Ser Ser
               245                 250                 255
Ala Leu Ile Gln His Gln Ile Thr His Thr Gly Glu Lys Pro Tyr
               260                 265                 270
Ile Cys Lys Glu Cys Gly Lys Ala Phe Thr Leu Ser Thr Ser Leu
               275                 280                 285
Tyr Lys His Leu Arg Thr His Thr Val Glu Lys Ser Tyr Arg Cys
               290                 295                 300
Lys Glu Cys Gly Lys Ser Phe Ser Arg Arg Ser Gly Leu Phe Ile
               305                 310                 315
His Gln Lys Ile His Ala Glu Glu Asn Pro Cys Lys Tyr Asn Pro
               320                 325                 330
Gly Arg Lys Ala Ser Ser Cys Ser Thr Ser Leu Ser Gly Cys Gln
               335                 340                 345
Arg Ile His Ser Arg Lys Lys Ser Tyr Leu Cys Asn Glu Cys Gly
               350                 355                 360
Asn Thr Phe Lys Ser Ser Ser Ser Leu Arg Tyr His Gln Arg Ile
               365                 370                 375
His Thr Gly Glu Lys Pro Phe Lys Cys Ser Glu Cys Gly Arg Ala
               380                 385                 390
Phe Ser Gln Ser Ala Ser Leu Ile Gln His Glu Arg Ile His Thr
               395                 400                 405
Gly Glu Lys Pro Tyr Arg Cys Asn Glu Cys Gly Lys Gly Phe Thr
               410                 415                 420
Ser Ile Ser Arg Leu Asn Arg His Arg Ile Ile His Thr Gly Glu
               425                 430                 435
Lys Phe Tyr Asn Cys Asn Glu Cys Gly Lys Ala Leu Ser Ser His
               440                 445                 450
Ser Thr Leu Ile Ile His Glu Arg Ile His Thr Gly Glu Lys Pro
               455                 460                 465
Cys Lys Cys Lys Val Cys Gly Lys Ala Phe Arg Gln Ser Ser Ala
```

```
                470                 475                 480
Leu Ile Gln His Gln Arg Met His Thr Gly Glu Arg Pro Tyr Lys
                485                 490                 495
Cys Asn Glu Cys Gly Lys Thr Phe Arg Cys Asn Ser Ser Leu Ser
                500                 505                 510
Asn His Gln Arg Ile His Thr Gly Glu Lys Pro Tyr Arg Cys Glu
                515                 520                 525
Glu Cys Gly Ile Ser Phe Gly Gln Ser Ser Ala Leu Ile Gln His
                530                 535                 540
Arg Arg Ile His Thr Gly Glu Lys Pro Phe Lys Cys Asn Thr Cys
                545                 550                 555
Gly Lys Thr Phe Arg Gln Ser Ser Ser Arg Ile Ala His Gln Arg
                560                 565                 570
Ile His Thr Gly Glu Lys Pro Tyr Glu Cys Asn Thr Cys Gly Lys
                575                 580                 585
Leu Phe Asn His Arg Ser Ser Leu Thr Asn His Tyr Lys Ile His
                590                 595                 600
Ile Glu Glu Asp Pro
                605
```

<210> 18
<211> 757
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 1907472

<400> 18

```
Met Gln Ser Ser Pro Asn Gly Gln Phe Val Ala Pro Ser Asp Ile
  1                 5                  10                  15
Gln Leu Lys Cys Asn Tyr Cys Lys Asn Ser Phe Cys Ser Lys Pro
                 20                  25                  30
Glu Ile Leu Glu Trp Glu Asn Lys Val His Gln Phe Cys Ser Lys
                 35                  40                  45
Thr Cys Ser Asp Asp Tyr Lys Lys Leu His Cys Ile Val Thr Tyr
                 50                  55                  60
Cys Glu Tyr Cys Gln Glu Glu Lys Thr Leu His Glu Thr Val Asn
                 65                  70                  75
Phe Ser Gly Val Lys Arg Pro Phe Cys Ser Glu Gly Cys Lys Leu
                 80                  85                  90
Leu Tyr Lys Gln Asp Phe Ala Arg Arg Leu Gly Leu Arg Cys Val
                 95                 100                 105
Thr Cys Asn Tyr Cys Ser Gln Leu Cys Lys Lys Gly Ala Thr Lys
                110                 115                 120
Glu Leu Asp Gly Val Val Arg Asp Phe Cys Ser Glu Asp Cys Cys
                125                 130                 135
Lys Lys Phe Gln Asp Trp Tyr Tyr Lys Ala Ala Arg Cys Asp Cys
                140                 145                 150
Cys Lys Ser Gln Gly Thr Leu Lys Glu Arg Val Gln Trp Arg Gly
                155                 160                 165
Glu Met Lys His Phe Cys Asp Gln His Cys Leu Leu Arg Phe Tyr
                170                 175                 180
```

```
Cys Gln Gln Asn Glu Pro Asn Met Thr Thr Gln Lys Gly Pro Glu
                185                 190                 195
Asn Leu His Tyr Asp Gln Gly Cys Gln Thr Ser Arg Thr Lys Met
                200                 205                 210
Thr Gly Ser Ala Pro Pro Pro Ser Pro Thr Pro Asn Lys Glu Met
                215                 220                 225
Lys Asn Lys Ala Val Leu Cys Lys Pro Leu Thr Met Thr Lys Ala
                230                 235                 240
Thr Tyr Cys Lys Pro His Met Gln Thr Lys Ser Cys Gln Thr Asp
                245                 250                 255
Asp Thr Trp Arg Thr Glu Tyr Val Pro Val Pro Ile Pro Val Pro
                260                 265                 270
Val Tyr Ile Pro Val Pro Met His Met Tyr Ser Gln Asn Ile Pro
                275                 280                 285
Val Pro Thr Thr Val Pro Val Pro Val Pro Val Pro Val Phe Leu
                290                 295                 300
Pro Ala Pro Leu Asp Ser Ser Glu Lys Ile Pro Ala Ala Ile Glu
                305                 310                 315
Glu Leu Lys Ser Lys Val Ser Ser Asp Ala Leu Asp Thr Glu Leu
                320                 325                 330
Leu Thr Met Thr Asp Met Met Ser Glu Asp Glu Gly Lys Thr Glu
                335                 340                 345
Thr Thr Asn Ile Asn Ser Val Ile Ile Glu Thr Asp Ile Ile Gly
                350                 355                 360
Ser Asp Leu Leu Lys Asn Ser Asp Pro Glu Thr Gln Ser Ser Met
                365                 370                 375
Pro Asp Val Pro Tyr Glu Pro Asp Leu Asp Ile Glu Ile Asp Phe
                380                 385                 390
Pro Arg Ala Ala Glu Glu Leu Asp Met Glu Asn Glu Phe Leu Leu
                395                 400                 405
Pro Pro Val Phe Gly Glu Glu Tyr Glu Glu Gln Pro Arg Pro Arg
                410                 415                 420
Ser Lys Lys Lys Gly Ala Lys Arg Lys Ala Val Ser Gly Tyr Gln
                425                 430                 435
Ser His Asp Asp Ser Ser Asp Asn Ser Glu Cys Ser Phe Pro Phe
                440                 445                 450
Lys Tyr Thr Tyr Gly Val Asn Ala Trp Lys His Trp Val Lys Thr
                455                 460                 465
Arg Gln Leu Asp Glu Asp Leu Leu Val Leu Asp Glu Leu Lys Ser
                470                 475                 480
Ser Lys Ser Val Lys Leu Lys Glu Asp Leu Leu Ser His Thr Thr
                485                 490                 495
Ala Glu Leu Asn Tyr Gly Leu Ala His Phe Val Asn Glu Ile Arg
                500                 505                 510
Arg Pro Asn Gly Glu Asn Tyr Ala Pro Asp Ser Ile Tyr Tyr Leu
                515                 520                 525
Cys Leu Gly Ile Gln Glu Tyr Leu Cys Gly Ser Asn Arg Lys Asp
                530                 535                 540
Asn Ile Phe Ile Asp Pro Gly Tyr Gln Thr Phe Glu Gln Glu Leu
                545                 550                 555
Asn Lys Ile Leu Arg Ser Trp Gln Pro Ser Ile Leu Pro Asp Gly
                560                 565                 570
Ser Ile Phe Ser Arg Val Glu Glu Asp Tyr Leu Trp Arg Ile Lys
                575                 580                 585
Gln Leu Gly Ser His Ser Pro Val Ala Leu Leu Asn Thr Leu Phe
                590                 595                 600
Tyr Phe Asn Thr Lys Tyr Phe Gly Leu Lys Thr Val Glu Gln His
```

```
                605                 610                 615
Leu Arg Leu Ser Phe Gly Thr Val Phe Arg His Trp Lys Lys Asn
                620                 625                 630
Pro Leu Thr Met Glu Asn Lys Ala Cys Leu Arg Tyr Gln Val Ser
                635                 640                 645
Ser Leu Cys Gly Thr Asp Asn Glu Asp Lys Ile Thr Thr Gly Lys
                650                 655                 660
Arg Lys His Glu Asp Asp Glu Pro Val Phe Glu Gln Ile Glu Asn
                665                 670                 675
Thr Ala Asn Pro Ser Arg Cys Pro Val Lys Met Phe Glu Cys Tyr
                680                 685                 690
Leu Ser Lys Ser Pro Gln Asn Leu Asn Gln Arg Met Asp Val Phe
                695                 700                 705
Tyr Leu Gln Pro Glu Cys Ser Ser Ser Thr Asp Ser Pro Val Trp
                710                 715                 720
Tyr Thr Ser Thr Ser Leu Asp Arg Asn Thr Leu Glu Asn Met Leu
                725                 730                 735
Val Arg Val Leu Leu Val Lys Asp Ile Tyr Asp Lys Asp Asn Tyr
                740                 745                 750
Glu Leu Asp Glu Asp Thr Asp
                755
```

<210> 19
<211> 154
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 1985458

<400> 19

```
Met Val Glu Lys Lys Thr Ser Val Arg Ser Gln Asp Pro Gly Gln
  1               5                  10                  15
Arg Arg Val Leu Asp Arg Ala Ala Arg Gln Arg Arg Ile Asn Arg
                 20                  25                  30
Gln Leu Glu Ala Leu Glu Asn Asp Asn Phe Gln Asp Asp Pro His
                 35                  40                  45
Ala Gly Leu Pro Gln Leu Gly Lys Arg Leu Pro Gln Phe Asp Asp
                 50                  55                  60
Asp Ala Asp Thr Gly Lys Lys Lys Lys Lys Thr Arg Gly Asp His
                 65                  70                  75
Phe Lys Leu Arg Phe Arg Lys Asn Phe Gln Ala Leu Leu Glu Glu
                 80                  85                  90
Gln Asn Leu Ser Val Ala Glu Gly Pro Asn Tyr Leu Thr Ala Cys
                 95                 100                 105
Ala Gly Pro Pro Ser Arg Pro Gln Arg Pro Phe Cys Ala Val Cys
                110                 115                 120
Gly Phe Pro Ser Pro Tyr Thr Cys Val Ser Cys Gly Ala Arg Tyr
                125                 130                 135
Cys Thr Val Arg Cys Leu Gly Thr His Gln Glu Thr Arg Cys Leu
                140                 145                 150
Lys Trp Thr Val
```

<210> 20
<211> 587
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 2726431

<400> 20
Met Asp Ser Val Val Phe Glu Asp Val Ala Val Asp Phe Thr Leu
1 5 10 15
Glu Glu Trp Ala Leu Leu Asp Ser Ala Gln Arg Asp Leu Tyr Arg
20 25 30
Asp Val Met Leu Glu Thr Phe Gln Asn Leu Ala Ser Val Gly Lys
35 40 45
Ile Trp Asp Ser Leu Ser Ile Glu Asp Gln Thr Thr Asn Gln Gly
50 55 60
Arg Asn Leu Ser Arg Asn His Gly Leu Glu Arg Leu Cys Glu Ser
65 70 75
Asn Asp Gln Cys Gly Glu Ala Leu Ser Gln Ile Pro His Leu Asn
80 85 90
Leu Tyr Lys Lys Ile Pro Pro Gly Val Lys Gln Tyr Glu Tyr Asn
95 100 105
Thr Tyr Gly Lys Val Phe Met His Arg Arg Thr Ser Leu Lys Ser
110 115 120
Pro Ile Thr Val His Thr Gly His Lys Pro Tyr Gln Cys Gln Glu
125 130 135
Cys Gly Gln Ala Tyr Ser Cys Arg Ser His Leu Arg Met His Val
140 145 150
Arg Thr His Asn Gly Glu Arg Pro Tyr Val Cys Lys Leu Cys Gly
155 160 165
Lys Thr Phe Pro Arg Thr Ser Ser Leu Asn Arg His Val Arg Ile
170 175 180
His Thr Ala Glu Lys Thr Tyr Glu Cys Lys Gln Cys Gly Lys Ala
185 190 195
Phe Ile Asp Phe Ser Ser Leu Thr Ser His Leu Arg Ser His Thr
200 205 210
Gly Glu Lys Pro Tyr Lys Cys Lys Glu Cys Gly Lys Ala Phe Ser
215 220 225
Tyr Ser Ser Thr Phe Arg Arg His Thr Ile Thr His Thr Gly Glu
230 235 240
Lys Pro Tyr Lys Cys Lys Glu Cys Ala Glu Ala Phe Ser Tyr Ser
245 250 255
Ser Thr Phe Arg Arg His Met Ile Ser His Thr Gly Glu Lys Pro
260 265 270
His Lys Cys Lys Glu Cys Gly Glu Ala Phe Ser Tyr Ser Ser Ala
275 280 285
Phe Arg Arg His Met Ile Thr His Thr Gly Glu Lys Pro Tyr Glu
290 295 300
Cys Lys Gln Cys Gly Lys Thr Phe Ile Tyr Leu Gln Ser Phe Arg
305 310 315
Arg His Glu Arg Ile His Thr Gly Glu Lys Pro Tyr Glu Cys Lys
320 325 330
Gln Cys Gly Lys Thr Phe Ile Tyr Pro Gln Ser Phe Arg Arg His

```
                335                 340                 345
Glu Arg Thr His Gly Gly Glu Lys Pro Tyr Glu Cys Asn Gln Cys
                350                 355                 360
Gly Lys Ala Phe Ser His Pro Ser Ser Phe Arg Gly His Met Arg
                365                 370                 375
Val His Thr Gly Glu Lys Pro Tyr Glu Cys Lys Gln Cys Gly Lys
                380                 385                 390
Thr Phe Asn Trp Pro Ile Ser Leu Arg Lys His Met Arg Thr His
                395                 400                 405
Thr Arg Glu Lys Pro Tyr Glu Cys Lys Gln Cys Gly Lys Ala Phe
                410                 415                 420
Ser Leu Ser Ala Cys Phe Arg Glu His Val Arg Met His Pro Glu
                425                 430                 435
Asp Lys Ser Tyr Glu Cys Lys Leu Cys Gly Lys Ala Phe Tyr Cys
                440                 445                 450
His Ile Ser Leu Gln Lys His Met Arg Arg His Thr Ala Glu Lys
                455                 460                 465
Leu Tyr Lys Cys Lys Gln Cys Gly Lys Ala Phe Ser Trp Pro Glu
                470                 475                 480
Leu Leu Gln Gln His Val Arg Thr His Thr Val Glu Lys Pro Tyr
                485                 490                 495
Glu Cys Lys Glu Cys Gly Lys Val Phe Lys Trp Pro Ser Ser Leu
                500                 505                 510
Pro Ile His Met Arg Leu His Thr Gly Glu Lys Pro Tyr Gln Cys
                515                 520                 525
Lys His Cys Gly Lys Ala Phe Asn Cys Ser Ser Ser Leu Arg Arg
                530                 535                 540
His Val Arg Ile His Thr Thr Glu Lys Gln Tyr Lys Cys Asn Val
                545                 550                 555
Gly His Pro Pro Ala Asn Glu Phe Met Cys Ser Ala Ser Glu Lys
                560                 565                 570
Ser His Gln Glu Arg Asp Leu Ile Lys Val Val Asn Met Val Leu
                575                 580                 585
Pro Leu
```

<210> 21
<211> 346
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 2743828

<400> 21

```
Met Ser Lys Pro Arg Ala Val Glu Ala Ala Ala Ala Ala Ala Ala
  1               5                  10                  15
Val Ala Ala Thr Ala Pro Gly Pro Glu Met Val Glu Arg Arg Gly
                20                  25                  30
Pro Gly Arg Pro Arg Thr Asp Gly Glu Asn Val Phe Thr Gly Gln
                35                  40                  45
Ser Lys Ile Tyr Ser Tyr Met Ser Pro Asn Lys Cys Ser Gly Met
                50                  55                  60
Arg Phe Pro Leu Gln Glu Glu Asn Ser Val Thr His His Glu Val
```

```
                  65                  70                  75
Lys Cys Gln Gly Lys Pro Leu Ala Gly Ile Tyr Arg Lys Arg Glu
                  80                  85                  90
Glu Lys Arg Asn Ala Gly Asn Ala Val Arg Ser Ala Met Lys Ser
                  95                 100                 105
Glu Glu Gln Lys Ile Lys Asp Ala Arg Lys Gly Pro Leu Val Pro
                 110                 115                 120
Phe Pro Asn Gln Lys Ser Glu Ala Ala Glu Pro Pro Lys Thr Pro
                 125                 130                 135
Pro Ser Ser Cys Asp Ser Thr Asn Ala Ala Ile Ala Lys Gln Ala
                 140                 145                 150
Leu Lys Lys Pro Ile Lys Gly Lys Gln Ala Pro Arg Lys Lys Ala
                 155                 160                 165
Gln Gly Lys Thr Gln Gln Asn Arg Lys Leu Thr Asp Phe Tyr Pro
                 170                 175                 180
Val Arg Arg Ser Ser Arg Lys Ser Lys Ala Glu Leu Gln Ser Glu
                 185                 190                 195
Glu Arg Lys Arg Ile Asp Glu Leu Ile Glu Ser Gly Lys Glu Glu
                 200                 205                 210
Gly Met Lys Ile Asp Leu Ile Asp Gly Lys Gly Arg Gly Val Ile
                 215                 220                 225
Ala Thr Lys Gln Phe Ser Arg Gly Asp Phe Val Val Glu Tyr His
                 230                 235                 240
Gly Asp Leu Ile Glu Ile Thr Asp Ala Lys Lys Arg Glu Ala Leu
                 245                 250                 255
Tyr Ala Gln Asp Pro Ser Thr Gly Cys Tyr Met Tyr Tyr Phe Gln
                 260                 265                 270
Tyr Leu Ser Lys Thr Tyr Cys Val Asp Ala Thr Arg Glu Thr Asn
                 275                 280                 285
Arg Leu Gly Arg Leu Ile Asn His Ser Lys Cys Gly Asn Cys Gln
                 290                 295                 300
Thr Lys Leu His Asp Ile Asp Gly Val Pro His Leu Ile Leu Ile
                 305                 310                 315
Ala Ser Arg Asp Ile Ala Ala Gly Glu Glu Leu Leu Tyr Asp Tyr
                 320                 325                 330
Gly Asp Arg Ser Lys Ala Ser Ile Glu Ala His Pro Trp Leu Lys
                 335                 340                 345
His
```

<210> 22
<211> 481
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 2998209

```
<400> 22
Met Asp Phe Gln Arg Ile Glu Leu Ala Gly Ala Val Gly Ser Lys
  1               5                  10                  15
Glu Glu Leu Glu Val Asp Phe Lys Lys Leu Lys Gln Ile Lys Asn
                 20                  25                  30
Arg Met Lys Lys Thr Asp Trp Leu Phe Leu Asn Ala Cys Val Gly
```

```
                35                  40                  45
Val Val Glu Gly Asp Leu Ala Ala Ile Glu Ala Tyr Lys Ser Ser
            50                  55                  60
Gly Gly Asp Ile Ala Arg Gln Leu Thr Ala Asp Glu Val Arg Leu
        65                  70                  75
Leu Asn Arg Pro Ser Ala Phe Asp Val Gly Tyr Thr Leu Val His
    80                  85                  90
Leu Ala Ile Arg Phe Gln Arg Gln Asp Met Leu Ala Ile Leu Leu
95                  100                 105
Thr Glu Val Ser Gln Gln Ala Ala Lys Cys Ile Pro Ala Met Val
            110                 115                 120
Cys Pro Glu Leu Thr Glu Gln Ile Arg Arg Glu Ile Ala Ala Ser
        125                 130                 135
Leu His Gln Arg Lys Gly Asp Phe Ala Cys Tyr Phe Leu Thr Asp
    140                 145                 150
Leu Val Thr Phe Thr Leu Pro Ala Asp Ile Glu Asp Leu Pro Pro
155                 160                 165
Thr Val Gln Glu Lys Leu Phe Asp Glu Val Leu Asp Arg Asp Val
            170                 175                 180
Gln Lys Glu Leu Glu Glu Glu Ser Pro Ile Ile Asn Trp Ser Leu
        185                 190                 195
Glu Leu Ala Thr Arg Leu Asp Ser Arg Leu Tyr Ala Leu Trp Asn
    200                 205                 210
Arg Thr Ala Gly Asp Cys Leu Leu Asp Ser Val Leu Gln Ala Thr
215                 220                 225
Trp Gly Ile Tyr Asp Lys Asp Ser Val Leu Arg Lys Ala Leu His
            230                 235                 240
Asp Ser Leu His Asp Cys Ser His Trp Phe Tyr Thr Arg Trp Lys
        245                 250                 255
Asp Trp Glu Ser Trp Tyr Ser Gln Ser Phe Gly Leu His Phe Ser
    260                 265                 270
Leu Arg Glu Glu Gln Trp Gln Glu Asp Trp Ala Phe Ile Leu Ser
275                 280                 285
Leu Ala Ser Gln Pro Gly Ala Ser Leu Glu Gln Thr His Ile Phe
            290                 295                 300
Val Leu Ala His Ile Leu Arg Arg Pro Ile Ile Val Tyr Gly Val
        305                 310                 315
Lys Tyr Tyr Lys Ser Phe Arg Gly Glu Thr Leu Gly Tyr Thr Arg
    320                 325                 330
Phe Gln Gly Val Tyr Leu Pro Leu Leu Trp Glu Gln Ser Phe Cys
335                 340                 345
Trp Lys Ser Pro Ile Ala Leu Gly Tyr Thr Arg Gly His Phe Ser
            350                 355                 360
Ala Leu Val Ala Met Glu Asn Asp Gly Tyr Gly Asn Arg Gly Ala
        365                 370                 375
Gly Ala Asn Leu Asn Thr Asp Asp Asp Val Thr Ile Thr Phe Leu
    380                 385                 390
Pro Leu Val Asp Ser Glu Arg Lys Leu Leu His Val His Phe Leu
395                 400                 405
Ser Ala Gln Glu Leu Gly Asn Glu Glu Gln Gln Glu Lys Leu Leu
            410                 415                 420
Arg Glu Trp Leu Asp Cys Cys Val Thr Glu Gly Gly Val Leu Val
        425                 430                 435
Ala Met Gln Lys Ser Ser Arg Arg Arg Asn His Pro Leu Val Thr
    440                 445                 450
Gln Met Val Glu Lys Trp Leu Asp Arg Tyr Arg Gln Ile Arg Pro
455                 460                 465
```

```
Cys Thr Ser Leu Ser Asp Gly Glu Glu Asp Glu Asp Asp Glu Asp
                470                 475                 480
Glu
```

<210> 23
<211> 179
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 3340296

<400> 23

```
Met Ser Thr Gly Ser Leu Ser Asp Val Glu Asp Leu Gln Glu Val
  1               5                  10                  15
Glu Met Leu Glu Cys Asp Gly Leu Lys Met Asp Ser Asn Lys Glu
             20                  25                  30
Phe Val Thr Ser Asn Glu Ser Thr Glu Glu Ser Ser Asn Cys Glu
             35                  40                  45
Asn Gly Ser Pro Gln Lys Gly Arg Gly Gly Leu Gly Lys Arg Arg
             50                  55                  60
Lys Ala Pro Thr Lys Lys Ser Pro Leu Ser Gly Val Ser Gln Glu
             65                  70                  75
Gly Lys Gln Val Gln Arg Asn Ala Ala Asn Ala Arg Glu Arg Ala
             80                  85                  90
Arg Met Arg Val Leu Ser Lys Ala Phe Ser Arg Leu Lys Thr Thr
             95                 100                 105
Leu Pro Trp Val Pro Pro Asp Thr Lys Leu Ser Lys Leu Asp Thr
            110                 115                 120
Leu Arg Leu Ala Ser Ser Tyr Ile Ala His Leu Arg Gln Ile Leu
            125                 130                 135
Ala Asn Asp Lys Tyr Glu Asn Gly Tyr Ile His Pro Val Asn Leu
            140                 145                 150
Thr Trp Pro Phe Met Val Ala Gly Lys Pro Glu Ser Asp Leu Lys
            155                 160                 165
Glu Val Val Thr Ala Ser Arg Leu Cys Gly Thr Thr Ala Ser
            170                 175
```

<210> 24
<211> 254
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 3536740

<400> 24

```
Met Ile Asp Glu Ile Leu Ser Lys Glu Thr Cys Asp Tyr Phe Glu
  1               5                  10                  15
```

```
Lys Leu Ser Leu Tyr Ser Val Cys Pro Ser Leu Val Val Arg Pro
                20                  25                  30
Lys Pro Leu His Ser Cys Thr Gly Ser Pro Ser Leu Arg Ala Tyr
                35                  40                  45
Pro Leu Leu Ser Val Ile Thr Arg Gln Pro Thr Val Ile Ser His
                50                  55                  60
Leu Val Pro Ala Thr Pro Gly Ile Ala Gln Ala Leu Ser Cys His
                65                  70                  75
Gln Val Thr Glu Ala Val Ser Ala Glu Ala Pro Gly Gly Glu Ala
                80                  85                  90
Leu Ala Ser Ser Glu Ser Glu Thr Glu Gln Pro Thr Pro Arg Gln
                95                  100                 105
Lys Lys Pro Arg Arg Ser Arg Thr Ile Phe Thr Glu Leu Gln Leu
                110                 115                 120
Met Gly Leu Glu Lys Lys Phe Gln Lys Gln Lys Tyr Leu Ser Thr
                125                 130                 135
Pro Asp Arg Leu Asp Leu Ala Gln Ser Leu Gly Leu Thr Gln Leu
                140                 145                 150
Gln Val Lys Thr Trp Tyr Gln Asn Arg Arg Met Lys Trp Lys Lys
                155                 160                 165
Met Val Leu Lys Gly Gly Gln Glu Ala Pro Thr Lys Pro Lys Gly
                170                 175                 180
Arg Pro Lys Lys Asn Ser Ile Pro Thr Ser Glu Glu Ile Glu Ala
                185                 190                 195
Glu Glu Lys Met Asn Ser Gln Ala Gln Gly Gln Glu Gln Leu Glu
                200                 205                 210
Pro Ser Gln Gly Gln Glu Glu Leu Cys Glu Ala Gln Glu Pro Lys
                215                 220                 225
Ala Arg Asp Val Pro Leu Glu Met Ala Glu Pro Pro Asp Pro Pro
                230                 235                 240
Gln Glu Leu Pro Ile Pro Ser Ser Glu Pro Pro Pro Leu Ser
                245                 250
```

<210> 25
<211> 498
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 082155

```
<400> 25
Met Asp Phe Ser Val Lys Val Asp Ile Glu Lys Glu Val Thr Cys
  1                 5                  10                  15
Pro Ile Cys Leu Glu Leu Leu Thr Glu Pro Leu Ser Leu Asp Cys
                20                  25                  30
Gly His Ser Phe Cys Gln Ala Cys Ile Thr Ala Lys Ile Lys Glu
                35                  40                  45
Ser Val Ile Ile Ser Arg Gly Glu Ser Ser Cys Pro Val Cys Gln
                50                  55                  60
Thr Arg Phe Gln Pro Gly Asn Leu Arg Pro Asn Arg His Leu Ala
                65                  70                  75
Asn Ile Val Glu Arg Val Lys Glu Val Lys Met Ser Pro Gln Glu
```

```
                    80                  85                  90
Gly Gln Lys Arg Asp Val Cys Glu His His Gly Lys Lys Leu Gln
                    95                 100                 105
Ile Phe Cys Lys Glu Asp Gly Lys Val Ile Cys Trp Val Cys Glu
                   110                 115                 120
Leu Ser Gln Glu His Gln Gly His Gln Thr Phe Arg Ile Asn Glu
                   125                 130                 135
Val Val Lys Glu Cys Gln Glu Lys Leu Gln Val Ala Leu Gln Arg
                   140                 145                 150
Leu Ile Lys Glu Asp Gln Glu Ala Glu Lys Leu Glu Asp Asp Ile
                   155                 160                 165
Arg Gln Glu Arg Thr Ala Trp Lys Asn Tyr Ile Gln Ile Glu Arg
                   170                 175                 180
Gln Lys Ile Leu Lys Gly Phe Asn Glu Met Arg Val Ile Leu Asp
                   185                 190                 195
Asn Glu Glu Gln Arg Glu Leu Gln Lys Leu Glu Glu Gly Glu Val
                   200                 205                 210
Asn Val Leu Asp Asn Leu Ala Ala Ala Thr Asp Gln Leu Val Gln
                   215                 220                 225
Gln Arg Gln Asp Ala Ser Thr Leu Ile Ser Asp Leu Gln Arg Arg
                   230                 235                 240
Leu Thr Gly Ser Ser Val Glu Met Leu Gln Asp Val Ile Asp Val
                   245                 250                 255
Met Lys Arg Ser Glu Ser Trp Thr Leu Lys Lys Pro Lys Ser Val
                   260                 265                 270
Ser Lys Lys Leu Lys Ser Val Phe Arg Val Pro Asp Leu Ser Gly
                   275                 280                 285
Met Leu Gln Val Leu Lys Glu Leu Thr Asp Val Gln Tyr Tyr Trp
                   290                 295                 300
Val Asp Val Met Leu Asn Pro Gly Ser Ala Thr Ser Asn Val Ala
                   305                 310                 315
Ile Ser Val Asp Gln Arg Gln Val Lys Thr Val Arg Thr Cys Thr
                   320                 325                 330
Phe Lys Asn Ser Asn Pro Cys Asp Phe Ser Ala Phe Gly Val Phe
                   335                 340                 345
Gly Cys Gln Tyr Phe Ser Ser Gly Lys Tyr Tyr Trp Glu Val Asp
                   350                 355                 360
Val Ser Gly Lys Ile Ala Trp Ile Leu Gly Val His Ser Lys Ile
                   365                 370                 375
Ser Ser Leu Asn Lys Arg Lys Ser Ser Gly Phe Ala Phe Asp Pro
                   380                 385                 390
Ser Val Asn Tyr Ser Lys Val Tyr Ser Arg Tyr Arg Pro Gln Tyr
                   395                 400                 405
Gly Tyr Trp Val Ile Gly Leu Gln Asn Thr Cys Glu Tyr Asn Ala
                   410                 415                 420
Phe Glu Asp Ser Ser Ser Ser Asp Pro Lys Val Leu Thr Leu Phe
                   425                 430                 435
Met Ala Val Pro Pro Cys Arg Ile Gly Val Phe Leu Asp Tyr Glu
                   440                 445                 450
Ala Gly Ile Val Ser Phe Phe Asn Val Thr Asn His Gly Ala Leu
                   455                 460                 465
Ile Tyr Lys Phe Ser Gly Cys Arg Phe Ser Arg Pro Ala Tyr Pro
                   470                 475                 480
Tyr Phe Asn Pro Trp Asn Cys Leu Val Pro Met Thr Val Cys Pro
                   485                 490                 495
Pro Ser Ser
```

<210> 26
<211> 1299
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 095477

<400> 26
Met Ala Ala Glu Thr Gln Thr Leu Asn Phe Gly Pro Glu Trp Leu
1 5 10 15
Arg Ala Leu Ser Ser Gly Gly Ser Ile Thr Ser Pro Pro Leu Ser
20 25 30
Pro Ala Leu Pro Lys Tyr Lys Leu Ala Asp Tyr Arg Tyr Gly Arg
35 40 45
Glu Glu Met Leu Ala Leu Phe Leu Lys Asp Asn Lys Ile Pro Ser
50 55 60
Asp Leu Leu Asp Lys Glu Phe Leu Pro Ile Leu Gln Glu Glu Pro
65 70 75
Leu Pro Pro Leu Ala Leu Val Pro Phe Thr Glu Glu Glu Gln Arg
80 85 90
Asn Phe Ser Met Ser Val Asn Ser Ala Ala Val Leu Arg Leu Thr
95 100 105
Gly Arg Gly Gly Gly Gly Thr Val Val Gly Ala Pro Arg Gly Arg
110 115 120
Ser Ser Ser Arg Gly Arg Gly Arg Gly Arg Gly Glu Cys Gly Phe
125 130 135
Tyr Gln Arg Ser Phe Asp Glu Val Glu Gly Val Phe Gly Arg Gly
140 145 150
Gly Gly Arg Glu Met His Arg Ser Gln Ser Trp Glu Glu Arg Gly
155 160 165
Asp Arg Arg Phe Glu Lys Pro Gly Arg Lys Asp Val Gly Arg Pro
170 175 180
Asn Phe Glu Glu Gly Gly Pro Thr Ser Val Gly Arg Lys His Glu
185 190 195
Phe Ile Arg Ser Glu Ser Glu Asn Trp Arg Ile Phe Arg Glu Glu
200 205 210
Gln Asn Gly Glu Asp Glu Asp Gly Gly Trp Arg Leu Ala Gly Ser
215 220 225
Arg Arg Asp Gly Glu Arg Trp Arg Pro His Ser Pro Asp Gly Pro
230 235 240
Arg Ser Ala Gly Trp Arg Glu His Met Glu Arg Arg Arg Arg Phe
245 250 255
Glu Phe Asp Phe Arg Asp Arg Asp Asp Glu Arg Gly Tyr Arg Arg
260 265 270
Val Arg Ser Gly Ser Gly Ser Ile Asp Asp Asp Arg Asp Ser Leu
275 280 285
Pro Glu Trp Cys Leu Glu Asp Ala Glu Glu Glu Met Gly Thr Phe
290 295 300
Asp Ser Ser Gly Ala Phe Leu Ser Leu Lys Lys Val Gln Lys Glu
305 310 315
Pro Ile Pro Glu Glu Gln Glu Met Asp Phe Arg Pro Val Asp Glu
320 325 330
Gly Glu Glu Cys Ser Asp Ser Glu Gly Ser His Asn Glu Glu Ala

```
                335                 340                 345
Lys Glu Pro Asp Lys Thr Asn Lys Lys Glu Gly Glu Lys Thr Asp
            350                 355                 360
Arg Val Gly Val Glu Ala Ser Glu Glu Thr Pro Gln Thr Ser Ser
            365                 370                 375
Ser Ser Ala Arg Pro Gly Thr Pro Ser Asp His Gln Ser Gln Glu
            380                 385                 390
Ala Ser Gln Phe Glu Arg Lys Asp Glu Pro Lys Thr Glu Gln Thr
            395                 400                 405
Glu Lys Ala Glu Glu Glu Thr Arg Met Glu Asn Ser Leu Pro Ala
            410                 415                 420
Lys Val Pro Ser Arg Gly Asp Glu Met Val Ala Asp Val Gln Gln
            425                 430                 435
Pro Leu Ser Gln Ile Pro Ser Asp Thr Ala Ser Pro Leu Leu Ile
            440                 445                 450
Leu Pro Pro Pro Val Pro Asn Pro Ser Pro Thr Leu Arg Pro Val
            455                 460                 465
Glu Thr Pro Val Val Gly Ala Pro Gly Met Gly Ser Val Ser Thr
            470                 475                 480
Glu Pro Asp Asp Glu Glu Gly Leu Lys His Leu Glu Gln Gln Ala
            485                 490                 495
Glu Lys Met Val Ala Tyr Leu Gln Asp Ser Ala Leu Asp Asp Glu
            500                 505                 510
Arg Leu Ala Ser Lys Leu Gln Glu His Arg Ala Lys Gly Val Ser
            515                 520                 525
Ile Pro Leu Met His Glu Ala Met Gln Lys Trp Tyr Tyr Lys Asp
            530                 535                 540
Pro Gln Gly Glu Ile Gln Gly Pro Phe Asn Asn Gln Glu Met Ala
            545                 550                 555
Glu Trp Phe Gln Ala Gly Tyr Phe Thr Met Ser Leu Leu Val Lys
            560                 565                 570
Arg Ala Cys Asp Glu Ser Phe Gln Pro Leu Gly Asp Ile Met Lys
            575                 580                 585
Met Trp Gly Arg Val Pro Phe Ser Pro Gly Pro Ala Pro Pro Pro
            590                 595                 600
His Met Gly Glu Leu Asp Gln Glu Arg Leu Thr Arg Gln Gln Glu
            605                 610                 615
Leu Thr Ala Leu Tyr Gln Met Gln His Leu Gln Tyr Gln Gln Phe
            620                 625                 630
Leu Ile Gln Gln Gln Tyr Ala Gln Val Leu Ala Gln Gln Gln Lys
            635                 640                 645
Ala Ala Leu Ser Ser Gln Gln Gln Gln Gln Leu Ala Leu Leu Leu
            650                 655                 660
Gln Gln Phe Gln Thr Leu Lys Met Arg Ile Ser Asp Gln Asn Ile
            665                 670                 675
Ile Pro Ser Val Thr Arg Ser Val Ser Val Pro Asp Thr Gly Ser
            680                 685                 690
Ile Trp Glu Leu Gln Pro Thr Ala Ser Gln Pro Thr Val Trp Glu
            695                 700                 705
Gly Gly Ser Val Trp Asp Leu Pro Leu Asp Thr Thr Thr Pro Gly
            710                 715                 720
Pro Ala Leu Glu Gln Leu Gln Gln Leu Glu Lys Ala Lys Ala Ala
            725                 730                 735
Lys Leu Glu Gln Glu Arg Arg Glu Ala Glu Met Arg Ala Lys Arg
            740                 745                 750
Glu Glu Glu Glu Arg Lys Arg Gln Glu Glu Leu Arg Arg Gln Gln
            755                 760                 765
```

Glu Glu Ile Leu Arg Arg Gln Gln Glu Glu Glu Arg Lys Arg Arg
770 775 780
Glu Glu Glu Glu Leu Ala Arg Arg Lys Gln Glu Glu Ala Leu Arg
785 790 795
Arg Gln Arg Glu Gln Glu Ile Ala Leu Arg Arg Gln Arg Glu Glu
800 805 810
Glu Glu Arg Gln Gln Gln Glu Glu Ala Leu Arg Arg Leu Glu Glu
815 820 825
Arg Arg Arg Glu Glu Glu Glu Arg Arg Lys Gln Glu Glu Leu Leu
830 835 840
Arg Lys Gln Glu Glu Glu Ala Ala Lys Trp Ala Arg Glu Glu Glu
845 850 855
Glu Ala Gln Arg Arg Leu Glu Glu Asn Arg Leu Arg Met Glu Glu
860 865 870
Glu Ala Ala Arg Leu Arg His Glu Glu Glu Glu Arg Lys Arg Lys
875 880 885
Glu Leu Glu Val Gln Arg Gln Lys Glu Leu Met Arg Gln Arg Gln
890 895 900
Gln Gln Gln Glu Ala Leu Arg Arg Leu Gln Gln Gln Gln Gln Gln
905 910 915
Gln Gln Leu Ala Gln Met Lys Leu Pro Ser Ser Ser Thr Trp Gly
920 925 930
Gln Gln Ser Asn Thr Thr Ala Cys Gln Ser Gln Ala Thr Leu Ser
935 940 945
Leu Ala Glu Ile Gln Lys Leu Glu Glu Glu Arg Glu Arg Gln Leu
950 955 960
Arg Glu Glu Gln Arg Arg Gln Gln Arg Glu Leu Met Lys Ala Leu
965 970 975
Gln Gln Gln Gln Gln Gln Gln Gln Gln Lys Leu Ser Gly Trp Gly
980 985 990
Asn Val Ser Lys Pro Ser Gly Thr Thr Lys Ser Leu Leu Glu Ile
995 1000 1005
Gln Gln Glu Glu Ala Arg Gln Met Gln Lys Gln Gln Gln Gln Gln
1010 1015 1020
Gln Gln His Gln Gln Pro Asn Arg Ala Arg Asn Asn Thr His Ser
1025 1030 1035
Asn Leu His Thr Ser Ile Gly Asn Ser Val Trp Gly Ser Ile Asn
1040 1045 1050
Thr Gly Pro Pro Asn Gln Trp Ala Ser Asp Leu Val Ser Ser Ile
1055 1060 1065
Trp Ser Asn Ala Asp Thr Lys Asn Ser Asn Met Gly Phe Trp Asp
1070 1075 1080
Asp Ala Val Lys Glu Val Gly Pro Arg Asn Ser Thr Asn Lys Asn
1085 1090 1095
Lys Asn Asn Ala Ser Leu Ser Lys Ser Val Gly Val Ser Asn Arg
1100 1105 1110
Gln Asn Lys Lys Val Glu Glu Glu Glu Lys Leu Leu Lys Leu Phe
1115 1120 1125
Gln Gly Val Asn Lys Ala Gln Asp Gly Phe Thr Gln Trp Cys Glu
1130 1135 1140
Gln Met Leu His Ala Leu Asn Thr Ala Asn Asn Leu Asp Val Pro
1145 1150 1155
Thr Phe Val Ser Phe Leu Lys Glu Val Glu Ser Pro Tyr Glu Val
1160 1165 1170
His Asp Tyr Ile Arg Ala Tyr Leu Gly Asp Thr Ser Glu Ala Lys
1175 1180 1185
Glu Phe Ala Lys Gln Phe Leu Glu Arg Arg Ala Lys Gln Lys Ala

```
                1190                1195                1200
Asn Gln Gln Arg Gln Gln Gln Gln Leu Pro Gln Gln Gln Gln Gln
                1205                1210                1215
Gln Pro Pro Gln Gln Pro Pro Gln Gln Pro Gln Gln Gln Asp Ser
                1220                1225                1230
Val Trp Gly Met Asn His Ser Thr Leu His Ser Val Phe Gln Thr
                1235                1240                1245
Asn Gln Ser Asn Asn Gln Gln Ser Asn Phe Glu Ala Val Gln Ser
                1250                1255                1260
Gly Lys Lys Lys Lys Lys Gln Lys Met Val Arg Ala Asp Pro Ser
                1265                1270                1275
Leu Leu Gly Phe Ser Val Asn Ala Ser Ser Glu Arg Leu Asn Met
                1280                1285                1290
Gly Glu Ile Glu Thr Leu Asp Asp Tyr
                1295
```

<210> 27
<211> 951
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 1399169

<400> 27

```
Met Ala Thr Gly Thr Gly Lys His Lys Leu Leu Ser Thr Gly Pro
  1               5                  10                  15
Thr Glu Pro Trp Ser Ile Arg Glu Lys Leu Cys Leu Ala Ser Ser
                20                  25                  30
Val Met Arg Ser Gly Asp Gln Asn Trp Val Ser Val Ser Arg Ala
                35                  40                  45
Ile Lys Pro Phe Ala Glu Pro Gly Arg Pro Pro Asp Trp Phe Ser
                50                  55                  60
Gln Lys His Cys Ala Ser Gln Tyr Ser Glu Leu Leu Glu Thr Thr
                65                  70                  75
Glu Thr Pro Lys Arg Lys Arg Gly Glu Lys Gly Glu Val Val Glu
                80                  85                  90
Thr Val Glu Asp Val Ile Val Arg Lys Leu Thr Ala Glu Arg Val
                95                 100                 105
Glu Glu Leu Lys Lys Val Ile Lys Glu Thr Gln Glu Arg Tyr Arg
               110                 115                 120
Arg Leu Lys Arg Asp Ala Glu Leu Ile Gln Ala Gly His Met Asp
               125                 130                 135
Ser Arg Leu Asp Glu Leu Cys Asn Asp Ile Ala Thr Lys Lys Lys
               140                 145                 150
Leu Glu Glu Glu Glu Ala Glu Val Lys Arg Lys Ala Thr Asp Ala
               155                 160                 165
Ala Tyr Gln Ala Arg Gln Ala Val Lys Thr Pro Pro Arg Arg Leu
               170                 175                 180
Pro Thr Val Met Val Arg Ser Pro Ile Asp Ser Ala Ser Pro Gly
               185                 190                 195
Gly Asp Tyr Pro Leu Gly Asp Leu Thr Pro Thr Thr Met Glu Glu
               200                 205                 210
```

```
Ala Thr Ser Gly Val Thr Pro Gly Thr Leu Pro Ser Thr Pro Val
                215                 220                 225
Thr Ser Phe Pro Gly Ile Pro Asp Thr Leu Pro Pro Gly Ser Ala
                230                 235                 240
Pro Leu Glu Ala Pro Met Thr Pro Val Thr Asp Asp Ser Pro Gln
                245                 250                 255
Lys Lys Met Leu Gly Gln Lys Ala Thr Pro Pro Pro Ser Pro Leu
                260                 265                 270
Leu Ser Glu Leu Leu Lys Lys Gly Ser Leu Leu Pro Thr Ser Pro
                275                 280                 285
Arg Leu Val Asn Glu Ser Glu Met Ala Val Ala Ser Gly His Leu
                290                 295                 300
Asn Ser Thr Gly Val Leu Leu Glu Val Gly Gly Val Leu Pro Met
                305                 310                 315
Ile His Gly Gly Glu Ile Gln Gln Thr Pro Asn Thr Val Ala Ala
                320                 325                 330
Ser Pro Ala Ala Ser Gly Ala Pro Thr Leu Ser Arg Leu Leu Glu
                335                 340                 345
Ala Gly Pro Thr Gln Phe Thr Thr Pro Leu Ala Ser Phe Thr Thr
                350                 355                 360
Val Ala Ser Glu Pro Pro Val Lys Leu Val Pro Pro Pro Val Glu
                365                 370                 375
Ser Val Ser Gln Ala Thr Ile Val Met Met Pro Ala Leu Pro Ala
                380                 385                 390
Pro Ser Ser Ala Pro Ala Val Ser Thr Thr Glu Ser Val Ala Pro
                395                 400                 405
Val Ser Gln Pro Asp Asn Cys Val Pro Met Glu Ala Val Gly Asp
                410                 415                 420
Pro His Thr Val Thr Val Ser Met Asp Ser Ser Glu Ile Ser Met
                425                 430                 435
Ile Ile Asn Ser Ile Lys Glu Glu Cys Phe Arg Ser Gly Val Ala
                440                 445                 450
Glu Ala Pro Val Gly Ser Lys Ala Pro Ser Ile Asp Gly Lys Glu
                455                 460                 465
Glu Leu Asp Leu Ala Glu Lys Met Asp Ile Ala Val Ser Tyr Thr
                470                 475                 480
Gly Glu Glu Leu Asp Phe Glu Thr Val Gly Asp Ile Ile Ala Ile
                485                 490                 495
Ile Glu Asp Lys Val Asp Asp His Pro Glu Val Leu Asp Val Ala
                500                 505                 510
Ala Val Glu Ala Ala Leu Ser Phe Cys Glu Glu Asn Asp Asp Pro
                515                 520                 525
Gln Ser Leu Pro Gly Pro Trp Glu His Pro Ile Gln Gln Glu Arg
                530                 535                 540
Asp Lys Pro Val Pro Leu Pro Ala Pro Glu Met Thr Val Lys Gln
                545                 550                 555
Glu Arg Leu Asp Phe Glu Glu Thr Glu Asn Lys Gly Ile His Glu
                560                 565                 570
Leu Val Asp Ile Arg Glu Pro Ser Ala Glu Ile Lys Val Glu Pro
                575                 580                 585
Ala Glu Pro Glu Pro Val Ile Ser Gly Ala Glu Ile Val Ala Gly
                590                 595                 600
Val Val Pro Ala Thr Ser Met Glu Pro Pro Glu Leu Arg Ser Gln
                605                 610                 615
Asp Leu Asp Glu Glu Leu Gly Ser Thr Ala Ala Gly Glu Ile Val
                620                 625                 630
Glu Ala Asp Val Ala Ile Gly Lys Gly Asp Glu Thr Pro Leu Thr
```

635 640 645
Asn Val Lys Thr Glu Ala Ser Pro Glu Ser Met Leu Ser Pro Ser
650 655 660
His Gly Ser Asn Pro Ile Glu Asp Pro Leu Glu Ala Glu Thr Gln
665 670 675
His Lys Phe Glu Met Ser Asp Ser Leu Lys Glu Glu Ser Gly Thr
680 685 690
Ile Phe Gly Ser Gln Ile Lys Asp Ala Pro Gly Glu Asp Glu Glu
695 700 705
Glu Asp Gly Val Ser Glu Ala Ala Ser Leu Glu Glu Pro Lys Glu
710 715 720
Glu Asp Gln Gly Glu Gly Tyr Leu Ser Glu Met Asp Asn Glu Pro
725 730 735
Pro Val Ser Glu Ser Asp Asp Gly Phe Ser Ile His Asn Ala Thr
740 745 750
Leu Gln Ser His Thr Leu Ala Asp Ser Ile Pro Ser Ser Pro Ala
755 760 765
Ser Ser Gln Phe Ser Val Cys Ser Glu Asp Gln Glu Ala Ile Gln
770 775 780
Ala Gln Lys Ile Trp Lys Lys Ala Ile Met Leu Val Trp Arg Ala
785 790 795
Ala Ala Asn His Arg Tyr Ala Asn Val Phe Leu Gln Pro Val Thr
800 805 810
Asp Asp Ile Ala Pro Gly Tyr His Ser Ile Val Gln Arg Pro Met
815 820 825
Asp Leu Ser Thr Ile Lys Lys Asn Ile Glu Asn Gly Leu Ile Arg
830 835 840
Ser Thr Ala Glu Phe Gln Arg Asp Ile Met Leu Met Phe Gln Asn
845 850 855
Ala Val Met Tyr Asn Ser Ser Asp His Asp Val Tyr His Met Ala
860 865 870
Val Glu Met Gln Arg Asp Val Leu Glu Gln Ile Gln Gln Phe Leu
875 880 885
Ala Thr Gln Leu Ile Met Gln Thr Ser Glu Ser Gly Ile Ser Ala
890 895 900
Lys Ser Leu Arg Gly Arg Asp Ser Thr Arg Lys Gln Asp Ala Ser
905 910 915
Glu Lys Asp Ser Val Pro Met Gly Ser Pro Ala Phe Leu Leu Ser
920 925 930
Leu Phe Asp Gly Gly Thr Arg Gly Arg Arg Cys Ala Ile Glu Ala
935 940 945
Asp Met Lys Met Lys Lys
950

<210> 28
<211> 282
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 1442069

<400> 28

```
Met Pro Lys Arg Lys Ala Ala Gly Gln Gly Asp Met Arg Gln Glu
  1               5                  10                  15
Pro Lys Arg Arg Ser Ala Arg Leu Ser Ala Met Leu Val Pro Val
             20                  25                  30
Thr Pro Glu Val Lys Pro Lys Arg Thr Ser Ser Ser Arg Lys Met
             35                  40                  45
Lys Thr Lys Ser Asp Met Met Glu Glu Asn Ile Asp Thr Ser Ala
             50                  55                  60
Gln Ala Val Ala Glu Thr Lys Gln Glu Ala Val Val Glu Glu Asp
             65                  70                  75
Tyr Asn Glu Asn Ala Lys Asn Gly Glu Ala Lys Ile Thr Glu Ala
             80                  85                  90
Pro Ala Ser Glu Lys Glu Ile Val Glu Val Lys Glu Glu Asn Ile
             95                 100                 105
Glu Asp Ala Thr Glu Lys Gly Gly Glu Lys Lys Glu Ala Val Ala
            110                 115                 120
Ala Glu Val Lys Asn Glu Glu Glu Asp Gln Lys Glu Asp Glu Glu
            125                 130                 135
Asp Gln Asn Glu Glu Lys Gly Glu Ala Gly Lys Glu Asp Lys Asp
            140                 145                 150
Glu Lys Gly Glu Glu Asp Gly Lys Glu Asp Lys Asn Gly Asn Glu
            155                 160                 165
Lys Gly Glu Asp Ala Lys Glu Lys Glu Asp Gly Lys Lys Gly Glu
            170                 175                 180
Asp Gly Lys Gly Asn Gly Glu Asp Gly Lys Glu Lys Gly Glu Asp
            185                 190                 195
Glu Lys Glu Glu Glu Asp Arg Lys Glu Thr Gly Asp Gly Lys Glu
            200                 205                 210
Asn Glu Asp Gly Lys Glu Lys Gly Asp Lys Lys Glu Gly Lys Asp
            215                 220                 225
Val Lys Val Lys Glu Asp Glu Lys Glu Arg Glu Asp Gly Lys Glu
            230                 235                 240
Asp Glu Gly Gly Asn Glu Glu Glu Ala Gly Lys Glu Lys Glu Asp
            245                 250                 255
Leu Lys Glu Glu Glu Glu Gly Lys Glu Glu Asp Glu Ile Lys Glu
            260                 265                 270
Asp Asp Gly Lys Lys Glu Glu Pro Gln Ser Ile Val
            275                 280
```

<210> 29
<211> 186
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 1596668

<400> 29

```
Met Asp Ala Asp Ser Asp Val Ala Leu Asp Ile Leu Ile Thr Asn
  1               5                  10                  15
Val Val Cys Val Phe Arg Thr Arg Cys His Leu Asn Leu Arg Lys
             20                  25                  30
Ile Ala Leu Glu Gly Ala Asn Val Ile Tyr Lys Arg Asp Val Gly
```

```
                  35                  40                  45
Lys Val Leu Met Lys Leu Arg Lys Pro Arg Ile Thr Ala Thr Ile
                  50                  55                  60
Trp Ser Ser Gly Lys Ile Ile Cys Thr Gly Ala Thr Ser Glu Glu
                  65                  70                  75
Glu Ala Lys Phe Gly Ala Arg Arg Leu Ala Arg Ser Leu Gln Lys
                  80                  85                  90
Leu Gly Phe Gln Val Ile Phe Thr Asp Phe Lys Val Val Asn Val
                  95                 100                 105
Leu Ala Val Cys Asn Met Pro Phe Glu Ile Arg Leu Pro Glu Phe
                 110                 115                 120
Thr Lys Asn Asn Arg Pro His Ala Ser Tyr Glu Pro Glu Leu His
                 125                 130                 135
Pro Ala Val Cys Tyr Arg Ile Lys Ser Leu Arg Ala Thr Leu Gln
                 140                 145                 150
Ile Phe Ser Thr Gly Ser Ile Thr Val Thr Gly Pro Asn Val Lys
                 155                 160                 165
Ala Val Ala Thr Ala Val Glu Gln Ile Tyr Pro Phe Val Phe Glu
                 170                 175                 180
Ser Arg Lys Glu Ile Leu
                 185
```

<210> 30
<211> 917
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 1977214

<400> 30

```
Met Ala Glu Thr Leu Ser Gly Leu Gly Asp Ser Gly Ala Ala Gly
  1               5                  10                  15
Ala Ala Ala Leu Ser Ser Ala Ser Ser Glu Thr Gly Thr Arg Arg
                 20                  25                  30
Leu Ser Asp Leu Arg Val Ile Asp Leu Arg Ala Glu Leu Arg Lys
                 35                  40                  45
Arg Asn Val Asp Ser Ser Gly Asn Lys Ser Val Leu Met Glu Arg
                 50                  55                  60
Leu Lys Lys Ala Ile Glu Asp Glu Gly Gly Asn Pro Asp Glu Ile
                 65                  70                  75
Glu Ile Thr Ser Glu Gly Asn Lys Lys Thr Ser Lys Arg Ser Ser
                 80                  85                  90
Lys Gly Arg Lys Pro Glu Glu Glu Gly Val Glu Asp Asn Gly Leu
                 95                 100                 105
Glu Glu Asn Ser Gly Asp Gly Gln Glu Asp Val Glu Thr Ser Leu
                110                 115                 120
Glu Asn Leu Gln Asp Ile Asp Ile Met Asp Ile Ser Val Leu Asp
                125                 130                 135
Glu Ala Glu Ile Asp Asn Gly Ser Val Ala Asp Cys Val Glu Asp
                140                 145                 150
Asp Asp Ala Asp Asn Leu Gln Glu Ser Leu Ser Asp Ser Arg Glu
                155                 160                 165
```

```
Leu Val Glu Gly Glu Met Lys Glu Leu Pro Glu Gln Leu Gln Glu
                170                 175                 180
His Ala Ile Glu Asp Lys Glu Thr Ile Asn Asn Leu Asp Thr Ser
                185                 190                 195
Ser Ser Asp Phe Thr Ile Leu Gln Glu Ile Glu Glu Pro Ser Leu
                200                 205                 210
Glu Pro Glu Asn Glu Lys Ile Leu Asp Ile Leu Gly Glu Thr Cys
                215                 220                 225
Lys Ser Glu Pro Val Lys Glu Glu Ser Ser Glu Leu Glu Gln Pro
                230                 235                 240
Phe Ala Gln Asp Thr Ser Ser Val Gly Pro Asp Arg Lys Leu Ala
                245                 250                 255
Glu Glu Glu Asp Leu Phe Asp Ser Ala His Pro Glu Glu Gly Asp
                260                 265                 270
Leu Asp Leu Ala Ser Glu Ser Thr Ala His Ala Gln Ser Ser Lys
                275                 280                 285
Ala Asp Ser Leu Leu Ala Val Val Lys Arg Glu Pro Ala Glu Gln
                290                 295                 300
Pro Gly Asp Gly Glu Arg Thr Asp Cys Glu Pro Val Gly Leu Glu
                305                 310                 315
Pro Ala Val Glu Gln Ser Ser Ala Ala Ser Glu Leu Ala Glu Ala
                320                 325                 330
Ser Ser Glu Glu Leu Ala Glu Ala Pro Thr Glu Ala Pro Ser Pro
                335                 340                 345
Glu Ala Arg Asp Ser Lys Glu Asp Gly Arg Lys Phe Asp Phe Asp
                350                 355                 360
Ala Cys Asn Glu Val Pro Pro Ala Pro Lys Glu Ser Ser Thr Ser
                365                 370                 375
Glu Gly Ala Asp Gln Lys Met Ser Ser Pro Glu Asp Asp Ser Asp
                380                 385                 390
Thr Lys Arg Leu Ser Lys Glu Glu Lys Gly Arg Ser Ser Cys Gly
                395                 400                 405
Arg Asn Phe Trp Val Ser Gly Leu Ser Ser Thr Thr Arg Ala Thr
                410                 415                 420
Asp Leu Lys Asn Leu Phe Ser Lys Tyr Gly Lys Val Val Gly Ala
                425                 430                 435
Lys Val Val Thr Asn Ala Arg Ser Pro Gly Ala Arg Cys Tyr Gly
                440                 445                 450
Phe Val Thr Met Ser Thr Ala Glu Glu Ala Thr Lys Cys Ile Asn
                455                 460                 465
His Leu His Lys Thr Glu Leu His Gly Lys Met Ile Ser Val Glu
                470                 475                 480
Lys Ala Lys Asn Glu Pro Val Gly Lys Lys Thr Ser Asp Lys Arg
                485                 490                 495
Asp Ser Asp Gly Lys Lys Glu Lys Ser Ser Asn Ser Asp Arg Ser
                500                 505                 510
Thr Asn Leu Lys Arg Asp Asp Lys Cys Asp Arg Lys Asp Asp Ala
                515                 520                 525
Lys Lys Gly Asp Asp Gly Ser Gly Glu Lys Ser Lys Asp Gln Asp
                530                 535                 540
Asp Gln Lys Pro Gly Pro Ser Glu Arg Ser Arg Ala Thr Lys Ser
                545                 550                 555
Gly Ser Arg Gly Thr Glu Arg Thr Val Val Met Asp Lys Ser Lys
                560                 565                 570
Gly Val Pro Val Ile Ser Val Lys Thr Ser Gly Ser Lys Glu Arg
                575                 580                 585
Ala Ser Lys Ser Gln Asp Arg Lys Ser Ala Ser Arg Glu Lys Arg
```

```
                590                 595                 600
Ser Val Val Ser Phe Asp Lys Val Lys Glu Pro Arg Lys Ser Arg
                605                 610                 615
Asp Ser Glu Ser His Ser Arg Val Arg Glu Arg Ser Glu Arg Glu
                620                 625                 630
Gln Arg Met Gln Ala Gln Trp Glu Arg Glu Glu Arg Glu Arg Leu
                635                 640                 645
Glu Ile Ala Arg Glu Arg Leu Ala Phe Gln Arg Gln Arg Leu Glu
                650                 655                 660
Arg Glu Arg Met Glu Arg Glu Arg Leu Glu Arg Glu Arg Met His
                665                 670                 675
Val Glu His Asp Gly Arg Arg Glu Gln Glu Arg Ile His Arg Glu
                680                 685                 690
Arg Glu Glu Leu Arg Arg Gln Gln Glu Leu Arg Tyr Glu Gln Glu
                695                 700                 705
Arg Arg Pro Ala Val Arg Arg Pro Tyr Asp Leu Asp Arg Arg Asp
                710                 715                 720
Asp Ala Tyr Trp Pro Glu Ala Lys Arg Ala Ala Leu Asp Glu Arg
                725                 730                 735
Tyr His Ser Asp Phe Asn Arg Gln Asp Arg Phe His Asp Phe Asp
                740                 745                 750
His Arg Asp Arg Gly Arg Tyr Pro Asp His Ser Val Asp Arg Arg
                755                 760                 765
Glu Gly Ser Arg Ser Met Met Gly Glu Arg Glu Gly Gln His Tyr
                770                 775                 780
Pro Glu Arg His Gly Gly Pro Glu Arg His Gly Gly Ala Ser Arg
                785                 790                 795
Asp Gly Trp Gly Gly Tyr Gly Ser Asp Lys Arg Met Ser Glu Gly
                800                 805                 810
Arg Gly Leu Pro Pro Pro Pro Arg Gly Arg Arg Asp Trp Gly Asp
                815                 820                 825
His Gly Arg Arg Glu Asp Asp Arg Ser Trp Gln Gly Thr Ala Asp
                830                 835                 840
Gly Gly Met Met Asp Arg Asp His Lys Arg Trp Gln Gly Gly Glu
                845                 850                 855
Arg Ser Met Ser Gly His Ser Gly Pro Gly His Met Met Asn Arg
                860                 865                 870
Gly Gly Met Ser Gly Arg Gly Ser Phe Ala Pro Gly Gly Ala Ser
                875                 880                 885
Arg Gly His Pro Ile Pro His Gly Gly Met Gln Gly Gly Phe Gly
                890                 895                 900
Gly Gln Ser Arg Gly Ser Arg Pro Ser Asp Ala Arg Phe Thr Arg
                905                 910                 915
Arg Tyr
```

<210> 31
<211> 392
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 2181282

<400> 31

```
Met Pro Lys Pro Ile Met Val Ile Pro Thr Leu Ala Ser Leu Ala
  1               5                  10                  15
Ser Pro Thr Thr Leu Gln Ser Gln Met Leu Gly Gly Leu Gly Gln
             20                  25                  30
Asp Val Leu Leu Asn Asn Ser Leu Thr Pro Lys Tyr Leu Gly Cys
             35                  40                  45
Lys Gln Asp Asn Ser Ser Ser Pro Lys Pro Ser Ser Val Phe Arg
             50                  55                  60
Asn Gly Phe Ser Gly Ile Lys Lys Pro Trp His Arg Cys His Val
             65                  70                  75
Cys Asn His His Phe Gln Phe Lys Gln His Leu Arg Asp His Met
             80                  85                  90
Asn Thr His Thr Asn Arg Arg Pro Tyr Ser Cys Arg Ile Cys Arg
             95                 100                 105
Lys Ser Tyr Val Arg Pro Gly Ser Leu Ser Thr His Met Lys Leu
            110                 115                 120
His His Gly Glu Asn Arg Leu Lys Lys Leu Met Cys Cys Glu Phe
            125                 130                 135
Cys Ala Lys Val Phe Gly His Ile Arg Val Tyr Phe Gly His Leu
            140                 145                 150
Lys Glu Val His Arg Val Val Ile Ser Thr Glu Pro Ala Pro Ser
            155                 160                 165
Glu Leu Gln Pro Gly Asp Ile Pro Lys Asn Arg Asp Met Ser Val
            170                 175                 180
Arg Gly Met Glu Gly Ser Leu Glu Arg Glu Asn Lys Ser Asn Leu
            185                 190                 195
Glu Glu Asp Phe Leu Leu Asn Gln Ala Asp Glu Val Lys Leu Gln
            200                 205                 210
Ile Lys Cys Gly Arg Cys Gln Ile Thr Ala Gln Ser Phe Ala Glu
            215                 220                 225
Ile Lys Phe His Leu Leu Asp Val His Gly Glu Glu Ile Glu Gly
            230                 235                 240
Arg Leu Gln Glu Gly Thr Phe Pro Gly Ser Lys Gly Thr Gln Glu
            245                 250                 255
Glu Leu Val Gln His Ala Ser Pro Asp Trp Lys Arg His Pro Glu
            260                 265                 270
Arg Gly Lys Pro Glu Lys Val His Ser Ser Ser Glu Glu Ser His
            275                 280                 285
Ala Cys Pro Arg Leu Lys Arg Gln Leu His Leu His Gln Asn Gly
            290                 295                 300
Val Glu Met Leu Met Glu Asn Glu Gly Pro Gln Ser Gly Thr Asn
            305                 310                 315
Lys Pro Arg Glu Thr Cys Gln Gly Pro Glu Cys Pro Gly Leu His
            320                 325                 330
Thr Phe Leu Leu Trp Ser His Ser Gly Phe Asn Cys Leu Leu Cys
            335                 340                 345
Ala Glu Met Leu Gly Arg Lys Glu Asp Leu Leu His His Trp Lys
            350                 355                 360
His Gln His Asn Cys Glu Asp Pro Ser Lys Leu Trp Ala Ile Leu
            365                 370                 375
Asn Thr Val Ser Asn Gln Gly Val Ile Glu Leu Ser Ser Glu Ala
            380                 385                 390
Glu Lys
```

<210> 32
<211> 1566
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 591290

<400> 32
tggttcagaa gcgaattctg caccgcgaac ctgagaggca ggtagctgcg gcgctgggct 60
ggcggcggcg agtccacgtg ctccccgcgg ccggttgaaa ccgttggcgg gcgctggctg 120
agaggcaatg tttgctgtct tccattggag tgactgaatt tctacatgac ggctttttga 180
caagacttaa aacctgtctt ggatagagaa tatttagcca tttacctaaa aatggtattt 240
tttacatgca atgcatgtgg tgaatcagtg aagaaaatac aagtggaaaa gcatgtgtct 300
gtttgcagaa actgtgaatg cctttcttgc attgactgcg gtaaagattt ctggggcgat 360
gactataaaa accacgtgaa atgcataagt gaagatcaga agtatggtgg caaaggctat 420
gaaggtaaaa cccacaaagg cgacatcaaa cagcaggcgt ggattcagaa aattagtgaa 480
ttaataaaga gacccaatgt cagccccaaa gtgagagaac ttttagagca aattagtgct 540
tttgacaacg ttcccaggaa aaaggcaaaa tttcagaatt ggatgaagaa cagtttaaaa 600
gttcataatg aatccattct ggaccaggtg tggaatatct tttctgaagc ttccaacagc 660
gaaccagtca ataaggaaca ggatcaacgg ccactccacc cagtggcaaa tccacatgca 720
gaaatctcca ccaaggttcc agcctccaaa gtgaaagacg ccgtggaaca gcaaggggag 780
gtgaagaaga ataaaagaga aaaaaaggaa gaacggcaga agaaaaggaa aagagaaaag 840
aaagaactaa agttagaaaa ccaccaggaa aactcaagga atcagaagcc taagaagcgc 900
aaaaagggac aggaggctga ccttgaggct ggtggggagg aagtccctga ggccaatggc 960
tctgcaggga agaggagcaa gaagaagaag cagcgcaagg acagcgccag tgaggaagag 1020
gcacgcgtgg gcgcagggaa gaggaagcgg aggcactcgg aagttgaaac agattctaag 1080
aagaaaaaga tgaagctccc agagcatcct gagggcggag aaccagaaga cgatgaggct 1140
cctgcaaaag gtaaattcaa ctggaaggga actattaaag caattctgaa acaggcccca 1200
gacaatgaaa taaccatcaa aaagctaagg aaaaaggttt tagctcagta ctacacagtg 1260
acagatgagc atcacagatc cgaagaggaa ctcctggtca tctttaacaa gaaaatcagc 1320
aagaacccta cctttaagtt attaaaggac aaagtcaagc ttgtgaaatg aacatttgtg 1380
tatttaaaaa ttgaatccat tctgctgact tcttcctttc actgctgttt ataaaatgtg 1440
taatgaattc taacaactca aattttgctt tttgaagctg tatttttaag ttaagaaaat 1500
atatttttgg tataactttt atgagaaaaa taaaatatat tctggtccaa acttctaaaa 1560
aaaaaa 1566

<210> 33
<211> 2338
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 815856

<400> 33
aggggaggtt ctccttttgc tattgtcatc acccagcaac accagattca ccgctcctgc 60
acagtcaaca tcatgtttgg aaccccgcaa gagcatcgca acatgcccca agcagatgcc 120
atggtgctgg tggccagaaa ttatgagcgt tacaagaatg agtgccggga gaaggaacgt 180
gaggagattg ccagacaggc agccaagatg gccgatgaag ccatcctgca ggaaagagag 240
agaggaggcc ctgaggaggg agtgcgtggg ggccaccctc cagccatcca gagcctcatc 300
aacctgctgg cagacaacag gtacctcact gctgaagaga ctgacaagat catcaactac 360
ctgcgagagc ggaaggagcg gctgatgagg agcagcaccg actctctgcc tggtgagcta 420
cgtggcaggg ccgaggcccg atttcccgcc aaccactcgg ggcgacctcg ggtgcctcgc 480

tgaagacaca gccaagctcc caaccgctcc agagcggcca agtgctcccc tctgctacac 540
ccactccatc tgcacccccc acctcccagc aagagcttca ggccaaaatc ctcagcctct 600
tcaatagtgg cacagtgacg gccaatagca gctctgcatc cccctcggtt gctgccggaa 660
acaccccaaa ccagaatttt tccacagcag caaacagcca gcctcaacaa agatcacagg 720
cttctggcaa tcagcctcca agcattttgg gacagggagg atctgctcag aacatgggcc 780
ccagacctgg ggctccttcc caagggcttt ttggccagcc ttccagtcgc ctggcacctg 840
ctagcaacat gactagccag aggcctgtgt cttccacagg tatcaacttt gacaatccaa 900
gtgtacagaa ggctctggat accctgatcc agagtggccc tgctctctcc cacctggtta 960
gccagaccac agcacagatg gggcagccac aggcccccat gggatcttac cagaggcatt 1020
actgaagcta aatctttcaa ctctccccag tcccctcatc ccctggcctc ctcccactta 1080
cttgttctaa atagagctgt ttggaggatg ttctctgcgc tcccaggccg gcatcgagtg 1140
tcatcaattt ctaccacctg ctctctcttc tgcccaaggc tgtgttgctt attccttaca 1200
aagtttatac tgcatttggg gctgtatctt tttttgtttt ttgttttgta gaaaataaat 1260
atctccgggg gcagtacagg tgtctgggct tgtatttgat ggggtttctc cggtccctgt 1320
ttctactgga tttggagcca ggcccagcta gccaagtttg gaatggcatt tgtcatgtca 1380
gtagccacca cctttgttca ttgtgaacct accaaggctt tccagcttca tacacattga 1440
ccagagctca agctcctgcc tgcaactcct gcctagagtt gaagaaaagc aaactggcct 1500
tggcaggcac agtgtcatca taccctcacc ccatatgttt ggggtctgct tgaggattca 1560
taaatcagcc actctggatt gttgaggaat ggccatggca gccacagaaa aaagaatttt 1620
tctctctgag ccaaggttgt tttttgtttt tttctctttt ctttttgtt ttcatttcat 1680
tggaagatct ccaatggact gaacagctcc agtcagcagc agttaccaca aactgtgaat 1740
ctgggcccca ccactcttcc ctgttaacca gttctgtcag catccccctc tccagcagca 1800
cttccatgaa gttggttctg agactctggc cgtgaacacc cgtttcttca gtgatttgtt 1860
ttgggctttt ggctcaaaac cccaggctct tgtttttgtc tagactctta ttctgtttcc 1920
tgagcagcag gaggtaggga ccactttgat gtcagacttc tggtagctgg acatgttctc 1980
gagatgggtg gctgttcgcg acttttgtac cagagtgaaa ttgttagaag gagggtttct 2040
ggctgtggtt ctaaatggag ccccaggaag ctgccctctc cccagggttt gtgctcagta 2100
gagccctgtg gatcacagtc ttgaggtcct ctagcagggg tgagggagag cagcgacttc 2160
agctgagtcc ctgccagtgg ttaagcaaac aatggtttca aaattcaagg tccccaaatg 2220
gcagcatttt atgttctgac ctgtttgtgt tatatagtgg ttttttttt cctctttgga 2280
actcttgtgt tgttaataaa atgaaatgat tacttttaa ttaaaatgaa aaaaaaaa 2338

<210> 34
<211> 870
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 996352

<400> 34
aaaaccgaca ctgaaagctg ccggacaaca aatcacaggt caaaattatg agttcttcgt 60
attatgtgaa cgcgcttttt agcaaatata cggcgggggc ttctctgttc caaaatgccg 120
agccgacttc ttgctccttt gctcccaact cacagagaag cggctacggg gcgggcgccg 180
gcgccttcgc ctcgaccgtt ccgggcttat acaatgtcaa cagccccctt tatcagagcc 240
cctttgcgtc cggctacggc ctgggcgccg acgcctacgg caacctgccc tgcgcctcct 300
acgaccaaaa catccccggg ctctgcagtg acctcgccaa aggcgcctgc gacaagacgg 360
acgagggcgc gctgcatggc gcggctgagg ccaatttccg catctacccc tggatgcggt 420
cttcaggacc tgacaggaag cggggccgcc agacctacac gcgctaccag acgctggagc 480
tggagaagga gttccacttc aaccgctacc tgatccggcg ccgccgcatt gaaatcgccc 540
acgcgctctg cctcaccgag cgccagatta agatctggtt ccagaaccgc cgcatgaagt 600
ggaagaaaga gcataaggac gaaggtccga ctgccgccgc agctcccgag ggcgccgtgc 660
cctctgccgc cgccactgct gccgcggaca aggccgacga ggaggacgat gatgaagaag 720
aggaagacga ggaggaatga ggggccgatc cggggccctc tctgcaccgg acagtcggaa 780
aagcgtcttt aagagactca ctggttttac ttacaaaaat gggaaaaata aaagaaaatg 840

taaaaaacaa aaacaaaaac aaaaaagcat 870

<210> 35
<211> 1365
<212> DNA
<213> Homo sapiens

<220>
<221> unsure
<222> 6, 61, 882, 883, 884, 885, 886, 887, 888, 889, 890, 891, 892, 893, 894, 895, <221> unsure
<222> 896, 897, 898, 899, 900, 901, 902, 903, 904, 905, 906, 907, 908, 912, 914, 916, <221> unsure
<222> 920, 922, 929, 932, 964, 967, 971, 975, 986, 1000, 1028, 1033, 1039, 1052, <221> unsure
<222> 1053, 1063, 1065, 1071, 1080, 1090, 1100, 1102, 1110, 1113, 1123, 1128, <221> unsure
<222> 1134, 1139, 1142, 1143, 1144, 1153, 1158, 1164, 1169, 1177, 1180, 1181, <221> unsure
<222> 1182, 1184, 1189, 1191, 1192, 1211, 1224, 1239, 1246, 1247, 1248, 1263, <221> unsure
<222> 1266, 1274, 1284, 1291, 1314, 1321, 1322, 1339, 1341, 1343, 1347, 1351, <221> unsure
<222> 1353, 1359, 1362, 1363, 1364
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte Clone No.: 1273778

<400> 35
cttcangcct ggcggcggct tccaagctaa ggaacggttt ggggcagtgt cgttcccgga 60
ngtcggccgc cgttacccgc tcaccagcta cgcggcgcgt caggtccgcg gaggcgcggg 120
ctcggggcgc ctgcgggacg gtgaggccct gctgaggact ccggcaagtg tgggtcgcgg 180
cgacggcggg gctaaggccc tgggtccgcg cgcggtttga ccacggccgg ggccttgggc 240
atttcctggc cttcctgttg agccgtgtaa acgcggggtg atgacggcgc cgacctcttg 300
gcactgttgt gagagcgaag tgggcgcgag agcagacgcc agctacagtt ttttttgggt 360
tatgtcgtca tgaagccggc gctttcagtt gtgcaacctt gaacaaatgg gacactgccc 420
atctctaaga taagaacctg gaaaggggac tctgttggcc attggaaatt gcagaataat 480
gtctcaggtg acatttagtg atgtggctat agacttctct catgaagagt gggcatgcct 540
agattctgct cagagggact tatacaagga tgtgatggtc cagaattatg agaacctggt 600
ctctgtaggt ctttccgtaa ctaagccata tgtgatcatg ttattggagg atggaaaaga 660
gccctggatg atggagaaaa aactgtcaaa agcttaccca tttcctttat cacactctgt 720
tcctgcttct gtgaactttg gattctctgc tctatttgag cattgttcag aagtcactga 780
aatatttgag ttgtcagaac tatgtgtttt ctgggtgctt catttcttat ccaattctcc 840
taattccact gtagaagctt ttttcaagaa gtaaaaaaaa annnnnnnnn nnnnnnnnnn 900
nnnnnnnngg cngngnttgn cnttcttgnt anatttcaga ttgggaatca agatgggaaa 960
acanggnatt ntcancaaag aagganattt gtgatgaagn ttcaccccaa ccagtaacaa 1020
tggaaaangt tgnaaaacna agttttgaat tnncaaaatt ctnantaagg ntttggaatn 1080
tccagaatgn gacacgtttn gnaagagccn ggnattccaa agngtacnct ttcntgggnc 1140
gnnngaaaca aanttctncc tgtnggggna tttcccncen nnantatgnt nnccaggttg 1200
agaggaattt nttccaaaag ggtnccggtc ataaaaggng gggggnnnaa tgggggggaaa 1260
aanacntttg aaantcgaag aaanggggggg ncggccttcc aacccggggc caancccgta 1320
nnccttcccc ggtctgttnt ngngggnaag ntntttacng gnnng 1365

<210> 36
<211> 2402
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 1509715

<400> 36
agcgcgggac aaagggaagc gaagccggag ctgcgggcgc tttttctgcc cgcggtgtct 60
cagattcatt cttaaggaac tgagaactta atcttccaaa atgtcaaaaa gaccatctta 120
tgccccacct cccaccccag ctcctgcaac acaaatgccc agcacaccag ggtttgtggg 180
atacaatcca tacagtcatc tcgcctacaa caactacagg ctgggagggg aacccgggca 240
ccaacagccg ggtcacggca tcctctggta tcacgattcc aaaaccccca aagccaccag 300
ataagccgct gatgccctac atgaggtaca gcagaaaggt ctgggaccaa gtaaaggctt 360
ccaaccctga cctaaagttg tgggagattg gcaagattat tggtggcatg tggcgagatc 420
tcactgatga agaaaaacaa gaatatttaa acgaatacga agcagaaaag atagagtaca 480
atgaatctat gaaggcctat cataattccc ccgcgtacct tgcttacata aatgcaaaaa 540
gtcgtgcaga agctgcttta gaggaagaaa gtcgacagag acaatctcgc atggagaaag 600
gagaaccgta catgagcatt cagcctgctg aagatccaga tgattatgat gatggctttt 660
caatgaagca tacagccacc gcccgtttcc agagaaacca ccgcctcatc agtgaaattc 720
ttagtgagag tgtggtgcca gacgttcggt cagttgtcac aacagctaga atgcaggtcc 780
tcaaacggca ggtccagtcc ttaatggttc atcagcgaaa actagaagct gaacttcttc 840
aaatagagga acgacaccag gagaagaaga ggaaattcct ggaaagcaca gattcattta 900
acaatgaact taaaaggttg tgcggtctga aagtagaagt ggatatggag aaaattgcag 960
ctgagattgc acaggcagag gaacaggccc gcaaaaggca ggaggaaagg gagaaggagg 1020
ccgcagagca agctgagcgc agtcagagca gcatcgttcc tgaggaagaa caagcagcta 1080
acaaaggcga ggagaagaaa gacgacgaga acattccgat ggagacagag gagacacacc 1140
ttgaagaaac aacagagagc caacagaatg gtgaagaagg cacgtctact cctgaggaca 1200
aggagagtgg gcaggagggg gtcgacagta tggcagagga aggaaccagt gatagtaaca 1260
ctggctcgga gagcaacagt gcaacagtgg aggagccacc aacagatccc ataccagaag 1320
atgagaaaaa agaataagtg ttgccttgtt ttgtgtgttc taaatacttt ttttaatgaa 1380
aaaatgtttt ttggttttaa tggtgttacg tggtttgtgt attaattttt tttcttgtcc 1440
atatcacacc accaaaggct tttggaccat ttagcatcat gagcctaatg gctcagtcag 1500
tcacctttct taagtgttgt gaagatggct cttttctttg gatcttgttt ctagccctca 1560
actgctgaaa gcctcagaat ttagattaat tgagaaaaca cccacctctt ttagagaatt 1620
atcctttgat gctgcagaat ctactcttac aatgccttcc tacagctcac tggggtgctt 1680
accaaagcca tagctttaaa ccttcccagt ccccatcagt agcttcctga aagtctcctc 1740
tcttgtttac ttctgcaaag ggtagcttct taaaaacgtg atcatgtatg agtatgtatt 1800
tgttcactta ccctttttta cttttaatca atgtcagata ccaagagttg tgttaagctg 1860
agtgtagtgt gtaactaact acacttggat cttactgatc cagaaatagt ccccatagtt 1920
agagtagtta cttatgaagt ggttattaaa gtgaacacag cacatataca ttatctatac 1980
tgctttttgt tatgattaat actgggtatg ttctggtaaa tccatcctta ttgtatagaa 2040
aaaaaattac ttttttacca ggttttccaa agacagaata gatcacaaag ctcaaggaat 2100
ttaatattct tgtaatggac tagataattc aaactgatta gcccattcca gaagaaaaac 2160
agctgggaat taagttaatc cacttgaaat tgttttacaa taatcagaac atccaaacct 2220
caaggctcag gatcccatag accagagccc acctttttga taaacttagt aaagtcttgg 2280
agactagaag caagatagtt tgtgacacat aagcttccca aaaactagaa tagattttta 2340
ctgaatagtg gtatatctga tggtatatgt ttcttaaagg tccaaatgta ataaaaaaaa 2400
aaaaaaaaaa aaaaaaaa
2418

<210> 37
<211> 866
<212> DNA

<213> Homo sapiens

<220>
<221> unsure
<222> 684, 698, 740, 755, 757, 776, 796, 805, 811, 824, 850, 856
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte Clone No.: 1676367

<400> 37
gccggggtgc gcacttgggc ccccctggcc atggcggcga aggtggacct gagcacctcc 60
accgactgga aggaggcgaa atcctttctg aagggcctga gtgacaagca gcgggaggaa 120
cattacttct gcaaggactt tgtcaggctg aagaagatcc cgacatggaa ggagatggcg 180
aaaggggtgg ctgtgaaggt ggaggagccc aggtataaaa aggacaagca gctcaatgag 240
aaaatctccc tgctccgcag cgacatcacc aagctggagg tggacgccat cgtcaacgcc 300
gccaacagct ccctgctcgg aggcggtggc gtggacggct gcattcatcg ggccgccggc 360
cccctgctta ccgacgagtg ccggaccctg cagagctgta agactggcaa ggccaagatc 420
accggcggct atcggctccc ggccaagtac gtcatccaca cagtggggcc catcgcctac 480
ggggagccca gcgccagtca ggctgccgag ctccgcagct gctacctgag cagtctggac 540
ctgctgctgg agcaccggct ccgctcggtg gcgttccct gcatctccac cggcgtgttt 600
ggctaccct gtgaggcggc cgccgagatc gtgctggcca cgctgcgaga gtggcttggg 660
agcagcacaa gggaaccccg gggnaattta aatttccngg aacccgggta actttgcaag 720
gcggttaacc aagctttttn cccctaataa attgnanttc cgtaatttaa gaagcntttg 780
gggggtaaaa tcaatngggt caatnaagct nggttttccc cggnggttga aaaatttggt 840
taatcccggn ttaaanaaat ttccca 866

<210> 38
<211> 1651
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 1734119

<400> 38
ggaaggtggt ggcaggcacc ggcctgggca gcttccaggc ctggctggcc acgaccacgg 60
cccgaggggg agcccgccag gccacgccgc actgagccac agccccgggg gccgcctccc 120
gggggccccc tgaggcactg aggcaccgag actggttctc cccgagagac tcggaaggtg 180
gggaacgagg ggactgtgtt tggggaggtg gctttttcgt ctgctgttga ctgaacacta 240
cagcgccctg tggttccggg cttcgcacag ctgtcccagg gatgggtcgc ctgtgctgcc 300
ttcgcccgcc gccacaccgg gaccctgcac ggctgcttct ggcctcgaca gatgacaaaa 360
gaaacagccc caaaatacga ccactccaac cagcagttcc cgcctgcctg cccgccactg 420
tcaggcctgc cctggcctcc tcgtccgcag ggctgtctgc tggcttctgg gggcagaaga 480
gcggggagcc ccgtggaagg gtcaggggag accaggtcag ggcagctaca tttctggtga 540
tcagccccat ggggagacgg ggctggcggg ataccgcccc ccccggcttc cccacaccac 600
ttctgtctca cccggaagcg tccttttttt gtgccaggtg tctacctaag agggttggtg 660
ccagaagccc cccatggcga gtgctggggc ccggcggtgc cctgggggag cagatggggc 720
caccctggc agggccgcta caacttttc cagcagcgga gccctctggg gggcctgtgc 780
ttgtggcatc tctgagggcc cagattgcac aaggtgacct ggccgtggcc tgagggtgga 840
gtcgcccagc acgcaggccg gggcgctgcg gggctaagta ttaggccttc ccagggaggg 900
ggcgtgccaa gcatcccaga gccgggctgg gaccgccaaa acgtcgtggc ctggatcctc 960
tgggtctgag tgcctgatcc cctgcccccc aaaaaaagca gaggtaggtg ttgcaggccc 1020
agggcagggg tgcctgcccc aggagagtcc caggcagtgg ttctcgtgcc agtggcaccc 1080

aggggcaagg acagccaacc cccacccttg ccacgtgtgg ggccacgtgg gcatgtgggg 1140
tgtgtgtttt taccttggtg aatctcacct gccaacgatt tctcgtgagt gccgaccacc 1200
ttctccgacc atgttacgcc cgggcggcag cagcccccgg ccactgcaaa cccatgccct 1260
gggtccctcg gctcccccag ggaggcatcc ccgtgccaat gtcccccagt ggtggcagca 1320
gatcctgtgg ccggcctggc ggacgggacc cagtgatact tgtatattac acagtcctga 1380
tttcagacaa tttcaacctt aatctattta aaaaagaata ttctatacaa gctgttttta 1440
agccttttac catttgaaat gcatgtgttg tgcgcgttgg ggatgggagg aggggctgag 1500
gagcggctca gtgtcacctc ccacagccac cggccctgac ccttaatcca gacaccgatg 1560
gaagtcgact tttcatatct ttctcctgaa atgaactctg ttttaaattg gaataaattt 1620
tgttcctaaa aaaaaaaaaa aaaaaaaaaa a 1651

<210> 39
<211> 1032
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 1944813

<400> 39
tattaatgaa gagtgactgc atgcaaacga caatatgtca ggaaagaaaa aaagatccca 60
tagaaatgtt tcattctgga cagctggtaa aagtctgtgc cccaatggtt cgatattcaa 120
agttggcttt taggacacta gtaagaaaat atagttgtga tctgtgttac acaccaatga 180
ttgttgccgc tgattttgtc aaatctataa aagccagaga cagcgaattt accacaaatc 240
aaggtgattg cccattgatt gttcagtttg ctgctaacga tgcaagactt ttatctgatg 300
ctgctcgtat agtctgtcct tatgcgaatg gaatagacat taactgtggt tgccctcaga 360
ggtgggcaat ggcagaaggt tatggggctt gcttaataaa caagccagag cttgttcaag 420
acatggtgaa acaagtaaga aatcaagtgg aaacccctgg attttcagtt tctattaaaa 480
taaggatcca tgatgacctt aaaagaactg tagatctttg tcaaaaggct gaagcaacag 540
gagtttcatg gattacagtc catggaagaa ctgctgaaga aagacatcag ccagtgcact 600
atgattccat taaaataatt aaggaaaata tgtctatacc tgtaattgct aatggagaca 660
tcagaagctt aaaggaagca gaaaatgtgt ggcggattac tgggacagat ggtgtgatgg 720
ttgcaagagg actcttagca aacccggcca tgtttgctgg atatgaggaa accccactga 780
aatgcatctg ggactgggtt gacattgctc ttgaactcgg gactccttac atgtgtttcc 840
atcaacattt aatgtacatg atggaaaaga taacttcaag gcaggaaaaa agggtattta 900
atgctctgtc aagcacatca gcaatcatag attaccttac agaccattat ggcatttgac 960
tagacttccc aaataatttt aatatacact tttagaccca cagtgaaacc acagaaggtc 1020
atattttgta cc 1032

<210> 40
<211> 1797
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 2683322

<400> 40
ttgtagtggg gctggagcag agcctgccgc gaacccccgg agcccacgat ccctcgtgcc 60
atccctcgaa tccaccagca cgagcgtccc acccgcgcct gggaccatgg ccactgactc 120
atgggccctg gcggtggacg agcaggaagc tgcggctgag tcgttgagca acttgcatct 180
taaggaagag aaaatcaaac cagataccaa tggtgctgtt gtcaagacca atgccaatgc 240
agagaagaca gatgaagaag agaaagagga cagagctgcc cagtccttac tcaacaagct 300

gatcagaagc aaccttgttg ataacacaaa ccaagtggaa gtcctgcagc gggatccaaa 360
ctcccctctg tactcggtga agtcttttga agagcttcgg ctgaaaccac agcttctcca 420
gggagtctat gccatgggct tcaatcgacc ctccaagata caagagaacg cattacccat 480
gatgcttgct gaacccccac agaatctgat tgcccagtct cagtctggca ctggtaaaac 540
agctgccttt gtcttagcca tgctcagccg agtggagcca tcagacagat acccccagtg 600
tctgtgcctc tccccaacat atgagctggc gcttcaaaca ggaaaagtga ttgagcagat 660
gggcaaattt tacccagaac tgaagcttgc ctatgccgtt cgaggcaata aattggaaag 720
aggccagaag atcagtgagc agattgtcat tggcacccct gggactgtgc tggactggtg 780
ctccaagctc aagttcattg atcccaagaa aatcaaggtg tttgttctgg atgaggctga 840
tgtcatgata gccactcagg gccaccaaga tcagagcatc cgcatccaga ggatgctgcc 900
caggaactgc cagatgctgc ttttctccgc caccttgaa gactctgtgt ggaagtttgc 960
ccagaaagtg gtcccagacc caaacgttat caaactgaag cgtgaggaag agaccctgga 1020
caccatcaag cagtactatg tcctgtgcag cagcagagac gagaagttcc aggccttgtg 1080
taacctctac ggggccatca ccattgctca agccatgatc ttctgccata ctcgcaaaac 1140
agctagttgg ctggcagcag agctctcaaa agaaggccac caggtggctc tgctgagtgg 1200
ggagatgatg gtggagcaga gggctgcagt gattgagcgc ttccgagagg gcaaagagaa 1260
ggttttggtg accaccaacg tgtgtgcccg cggcattgat gttgaacaag tgtctgtcgt 1320
catcaacttt gatcttcccg tggacaagga cgggaatcct gacaatgaga cctacctgca 1380
ccggatcggg cgcacgggcc gctttggcaa gaggggcctg gcagtgaaca tggtggacag 1440
caagcacagc atgaacatcc tgaacagaat ccaggagcat tttaataaga agatagaaag 1500
attggacaca gatgatttgg acgagattga gaaaatagcc aactgagaag ctccaccagc 1560
cactgatgcc agccctggca ctgcccctgc acaggagaca agtgcgttca gggcacaggc 1620
cccgacatca ccccaaggac aacggcacaa gtagagagaa actacctacc tcacttcaaa 1680
ttatgtttgg acttgacaaa aatgtatgca aatgatgggg gatggtagaa aaaaattatt 1740
tacacaacct tggaagatta ggcatgaata cacagagatt tacctttaaa aaaaaaa 1797

<210> 41
<211> 1987
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 2684552

<400> 41
tgacagagga ggctccgtgt ctgcagctag tgtgtcaact cagcgtttct cctctcgtcc 60
ctggaagagc taaagatggc tgaatttcta gatgaccagg aaactcgact gtgtgacaac 120
tgcaaaaaag aaattcctgt gtttaacttt accatccatg agatccactg tcaaaggaac 180
attggtatgt gtcctacctg taaggaacca tttcccaaat ctgacatgga gactcacatg 240
gctgcagaac actgtcaggt gacctgcaaa tgtaacaaga agttggagaa gaggctgtta 300
aagaagcatg aggagactga gtgcccttg cggcttgctg tctgccagca ctgtgattta 360
gaactttcca ttctcaaact gaaggaacat gaagattatt gtggtgcccg gacggaacta 420
tgtggcaact gtggtcgcaa tgtccttgtg aaagatctga agactcaccc tgaagtttgt 480
gggagagagg gggaggaaaa gagaaatgag gttgccatac ctcctaatgc atatgatgaa 540
tcttggggtc aggatggaat ctggattgca tcccaactcc tcagacaaat tgaagctctg 600
gacccaccca tgaggctgcc gcgaaggccc ctgagagcct ttgaatcaga tgttttccac 660
aatagaacta ccaaccaaag gaacattaca gcccaggttt caattcagaa taatctgttt 720
gaagaacaag agaggcagga aaggaataga ggccaacagc cccccaaaga gggtggtgaa 780
gagagtgcaa acttggactt catgttggcc ctaagtctgc aaaatgaagg ccaagcctcc 840
agtgtggcag agcaggactt ctggagggcc gtatgtgagg ccgaccagtc tcatggcggt 900
cccaggtctc tcagtgacat aaagggtgca gctgacgaga tcatgttgcc ttgtgaattt 960
tgtgaggagc tctacccaga ggaactgctg attgaccatc agacaagctg taacccttca 1020
cgtgccttac cttcactcaa tactggcagc tcttccccca gaggggtgga ggaacctgat 1080
gtcatcttcc agaacttctt gcaacaggct gcaagtaacc agttagactc tttgatgggc 1140
ctgagcaatt cacaccctgt ggaggagagc atcattatcc catgtgaatt ctgtggggta 1200

cagctggaag aggaggtgct gttccatcac caggaccagt gtgaccaacg cccagccact 1260
gcaaccaacc atgtgacaga ggggattcct agactggatt cccagcctca agagacctca 1320
ccagagctgc ccaggaggcg tgtcagacac cagggagacc tgtcttctgg ttacctggat 1380
gatactaagc aggaaacagc taatgggccc acctcctgtc tgcctcccag ccgacccatt 1440
aacaatatga cagctaccta taaccagcta tcgagatcaa catcaggccc cagacctggg 1500
tgccagccca gctctccttg tgtgccgaag ctcagcaact cagacagcca ggacatccag 1560
gggcggaatc gagacagcca gaatggggcc atagcccctg ggcacgtttc agtgattcgc 1620
cctcctcaaa atctctaccc agaaaacatt gtgccctctt tctcccctgg gccttcaggg 1680
agatacggag ctagtggtag gagtgaaggt ggcaggaatt cccgggtcac ccctgcagct 1740
gccaactacc gcagcagaac tgcaaaggca aagccttcca agcaacaggg agctggggat 1800
gcagaagagg aagaggagga gtaatggtgt ctccagagac tttaacaatc gggtcccggc 1860
ttctttgtga ccggaagaac cttgtcgctg ttgcagggcc acctctcttg gccctttggg 1920
gggggagagt tttttccaaa gttataatat tttccaaagg tatggcccat tgtgggctct 1980
ttaaggg 1987

<210> 42
<211> 2295
<212> DNA
<213> Homo sapiens

<220>
<221> unsure
<222> 2228
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte Clone No.: 2830310

<400> 42
gcgggatagg taaacccgcc gcgcactggc gggacgctgg gccaacgagg tctggggtct 60
tggcgcggag ttcatgggcc gccggccagg cttgtgccag tggctttgga tccctggcct 120
gcttccccca acccagcctc cagaccagcg ctggcctgct ccacacactc cgcagcaccg 180
tcccgggaac cgagcgtcgg cggggaggcc gcctctccgc aataccccat gtttccccat 240
cactcgagga gtctgggcag agactggact acacgtggga gaatctgcaa aggtgttgct 300
ggaacagaca tatttctagt tgtatgaggt ggcctggaca ttactctcga gctccttacc 360
catacttcag tagtaggcat ttttcactaa attggagacc accctgtttg tttgagtcta 420
gaactcagtt ccagtactgt aactggagac ctgacaacct gagccagaca tctttgattc 480
atctctctag ttacgtcatg aacgctgagg gagatgagcc ttcatcaaaa cgaagaaaac 540
accaaggtgt gataaagcgg aattgggaat atatatgtag ccatgataaa gaaaaaacga 600
agatcctagg agacaaaaat gttgatccca aatgtgaaga cagtgagaac aagtttgact 660
tttcagtgat gtcctataat atactttcac aagatttact ggaagataac tctcaccttt 720
atagacattg ccggcggcca gtattacact ggagttttag gtttcccaat attctgaaag 780
aaattaaaca ttttgatgca gacgtacttt gtttgcaaga agttcaagaa gatcattatg 840
gagcagagat caggccaagt ttggaatcac tgggttatca ctgtgaatat aagatgcgga 900
caggaaggaa acctgatggc tgtgctattt gcttcaaaca ttccaaattt tcactcttgt 960
cagtgaaccc agtggaattc ttccgccctg atatttctct gttggacaga gacaatgttg 1020
gattagtttt actcttacag cccaaaattc catatgctgc ctgccctgca atctgcgtag 1080
caaatacgca tctgttgtat aatccaaggc gaggtgatat taagctgacg caattggcaa 1140
tgctactggc agagatttcc agtgttgccc accagaaaga tggcagcttc tgccctattg 1200
ttatgtgtgg tgactttaat tctgttcctg gttctccact atatagtttc ataaaggaag 1260
gaaaattgaa ttatgaagga cttcccatag gaaagacagt gatctgacac aaacacagct 1320
gaagcaaaca gaggtcctag tgacagctga aaaattgtct tcaaatttac agcaccattt 1380
cagtttgtca tctgtttatt cacattactt tcctgacact ggaattccag aagtgaccac 1440
ctgtcattcc cgaagtgcca taactgtgga ttatattttc tactctgcag aaaaggaaga 1500

tgttgctggg cacccaggag ctgaagttgc tttggttggt ggcttgaaac ttctagctag 1560
actgtcactt cttacagaac aagacttatg gactgttaat ggacttccaa acgaaaataa 1620
ctcttcagat catctgcctt tattggcaaa gttcagactt gagctctgac tctctttgat 1680
cacatactaa ttttctttcc aatttgtatt gtttttcaaa gaatgtaaag ttcttaagtg 1740
tatgcatgtt gtttatttttt gcactgtgga gattctgaag cggttatgtt agatgctttg 1800
aaactccata tcagaagaaa taactttata acaatttttt ttaataatga aaaatatttt 1860
cgtgacaagt gagctctaaa ttctctttat tgtaaaagag atgtaaaggt tttatattct 1920
aaatcctagt aaaattgaca gtgatttttа aatataatgc atcttccttt gtctgcttag 1980
taaaaaattt catttcataa ttttggcaag ctctgtagtg gatccaaagt atctttgagt 2040
tcttgcaaac tacaagttgt ttcctttcca gaaggcttga tttcattagg agacccctct 2100
attgagttct aaatagttta tcttagaaag ccttgggtca ttcacaggta tccaaccagc 2160
cattgtttag tttgtttttg aaggggtttg ataatgcttt ttaagttgta cagaatgctt 2220
aatccacnta ttactgtcct gagccagtaa tatgcctgca tcgtggtggg gaatgtttgg 2280
gaaatataag ccagc 2295

<210> 43
<211> 2819
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 2963346

<400> 43
taggacttca acatggcggc tgcggcactg gcggtggcta cggtgacggc ctggcccgga 60
gcgggcagag ttggaggtgg tggcgttcgc tctccctagg ggctgtcggg agctcagcgg 120
ggaccgagcc tgggaggccg gccggtgcca gcacctttcg gcttctgaga cggcggcagc 180
agcggcattc aggttctaaa tggcttctaa gaagttgggt gcagattttc atgggacttt 240
cagttacctt gatgatgtcc catttaagac aggagacaaa ttcaaaacac cagctaaagt 300
tggtctacct attggcttct ccttgcctga ttgtttgcag gttgtcagag aagtacagta 360
tgacttctct ttggaaaaga aaaccattga gtgggctgaa gagattaaga aaatcgaaga 420
agccgagcgg gaagcagagt gcaaaattgc ggaagcagaa gctaaagtga attctaagag 480
tggcccagag ggcgatagca aaatgagctt ctccaagact cacagtacag ccacaatgcc 540
acctcctatt aaccccatcc tcgccagctt gcagcacaac agcatcctca caccaactcg 600
ggtcagcagt agtgccacga aacagaaagt tctcagccca cctcacataa aggcggattt 660
caatcttgct gactttgagt gtgaagaaga cccatttgat aatctggagt taaaaactat 720
tgatgagaag gaagagctga gaaatattct ggtaggaacc actggaccca ttatggctca 780
gttattggac aataacttgc ccaggggagg ctctgggtct gtgttacagg atgaggaggt 840
cctggcatcc ttggaacggg caaccctaga tttcaagcct cttcataaac ccaatggctt 900
tataacctta ccacagttgg gcaactgtga aaagatgtca ctgtcttcca aagtgtccct 960
ccccctata cctgcagtaa gcaatatcaa atccctgtct ttccccaaac ttgactctga 1020
tgacagcaat cagaagacag ccaagctggc gagcactttc catagcacat cctgcctccg 1080
caatggcacg ttccagaatt ccctaaagcc ttccacccaa agcagtgcca gtgagctcaa 1140
tgggcatcac actcttgggc tttcagcttt gaacttggac agtggcacag agatgccagc 1200
cctgacatcc tcccagatgc cttccctctc tgttttgtct gtgtgcacag aggaatcatc 1260
acctccaaat actggtccca cggtcacccc tcctaatttc tcagtgtcac aagtgcccaa 1320
catgcccagc tgtccccagg cctattctga actgcagatg ctgtcccccа gcgagcggca 1380
gtgtgtggag acggtggtca acatgggcta ctcgtacgag tgtgtcctca gagccatgaa 1440
gaagaaagga gagaatattg agcagattct cgactatctc tttgcacatg gacagctttg 1500
tgagaagggc ttcgaccctc ttttagtgga agaggctctg gaaatgcacc agtgttcaga 1560
agaaaagatg atggagtttc ttcagttaat gagcaaattt aaggagatgg gctttgagct 1620
gaaagacatt aaggaagttt tgctattaca caacaatgac caggacaatg ctttggaaga 1680
cctcatggct cgggcaggag ccagctgaga ccaggccctg cctaggccct gccgcagaac 1740
caccatccct gggaggccct gcagagccca cctgtgggga aagagaaggg gcagcttccg 1800
gattttcttt tgggggttag aaggtcaggt gtggagactg ctcgccagtc tctgtgagcc 1860

taggccctga gctggggagg tggggaagat tcgggcatgt gagtgccccc agaactgtcc 1920
tggctccttc cgtattaaac gcatttgcat tttgagaagt gtccttccca cttcagccct 1980
ccggagagac taccctagtc tttctggggt gtttatgtcc tcagctgaag cctggcctag 2040
ttgctgagag gggctgggga gatggggcgg gagggccaga ctcagtgctg ctgtggagct 2100
aggtgcttcc cccttccccc gagactggtg gactgaactc cagtcaagtt gagttcaagt 2160
gaaagattct tccagggttt tatttttcc cctcctaaca aagtctcata gtgttaacac 2220
tggttctgca atatctctga ggtgcaaaga atgcactttt ccctatgggg cccagagttt 2280
gccttttctg ccaggcagtc accacgcttc cctaccccag cctgtttctt ttggcttggt 2340
ttggaccaca gtcctctgct acccagggtt ttagagcccc tgctctagga aacagtttaa 2400
gaaatcattg gccccttccc agcacattga atgggtaagc agacaggcca tgatttagtt 2460
ggccagcact aactccacct ctgttctcct tgaacagctt cccctccagc ccactgcttt 2520
aggatgacac aatgaataac acctagtcat agaaatcagt ctctctggtt tgttttgtat 2580
tatgttgtac atcattaaag atctaaatac aaaggatata cagtcttgaa tctaaaataa 2640
tttgctaact aactattttg attcttcaga gagaactact aataaaaatc taaaaggtaa 2700
aaaaaaaaaa aaaaaaaaaa gggcggcaga tctagaggaa ccaagcttaa gtaagcgagc 2760
atgcgacatc atagatcatc tatagtagtc acctaaattc aatttcagtg ccaacaaag 2819

<210> 44
<211> 1459
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 2994234

<400> 44
cttcttgaga ccgaggaccg aaatcccggc tccaggcctc ggggactgcg gactgtgggg 60
aggctggccg gagagagagg gaaggacggg gcctggcccc cgggactccc tgtgccttgc 120
ttggagctga cgccgacggt ttattgcagg gaactgacaa gatcacattt tgagaagaag 180
ttggaaagaa tcccaagtgg atgaactgaa tatctggatg aggacaagat ctgtggggag 240
agactgtaag atagaatgag tccatttaag tcccaggacg gtggaaacta gctagtagat 300
tgcagccatg ttgtggaagc tgctgctgag atcccagtcc tgcaggctgt gttctttcag 360
aaagatgcga tcacctccaa aatacagacc tttcttagca tgcttcacct atacaactga 420
taaacagtcg agcaaagaaa atacaagaac agtggaaaag ctctataaat gttcagttga 480
cattaggaaa attcgtagat taaaaggatg ggtacttta gaggatgaaa cctatgttga 540
agaaattgcg aatattttac aagaactagg tgccgatgag actgctgtag ccagtatttt 600
ggaacgctgc ccggaagcaa ttgtctgtag tccaaccgct gttaacaccc agagaaaact 660
ctggcagttg gtctgcaaaa atgaggaaga gttaatcaag ttaatagagc agtttccaga 720
atctttcttt actattaaag accaagagaa ccagaagctg aatgttcagt tctttcaaga 780
gttgggacta aaaaatgtgg tcattagcag acttttgaca gctgcaccta atgtttttca 840
taatcctgtt gagaagaata agcaaatggt aagaattctc caagagagtt atctagatgt 900
aggtggctct gaggccaaca tgaaagtttg gctactaaaa ttgttaagcc aaaacccatt 960
tattttgtta aattctccca cagctataaa ggaaacacta gaatttctcc aggagcaagg 1020
tttcaccagc tttgaaattc tccagcttct atccaaactc aaaggatttc tttttcaact 1080
ttgcccaaga agtatacaga atagtatttc cttctctaaa aatgctttta aatgcacaga 1140
tcatgacctg aagcaattag ttttgaaatg tcctgccctt ttatattatt ctgttccagt 1200
tttagaagag agaatgcaag gattattgag agaaggaatt tccatagctc agataagaga 1260
gacgccaatg gttcttgaat taacaccaca gatagtacag tacaggataa ggaaactgaa 1320
ttcctcaggc tacagaataa aggatggaca tctagcaaat ctaaatggat caaaaaaaga 1380
gtttgaagct aatttttggca aaattcaggc caaaaaaagt aaggccatta tttaaccctg 1440
tggcaccatt aaatgttga 1459

<210> 45
<211> 2733

<212> DNA
<213> Homo sapiens

<220>
<221> unsure
<222> 61, 63, 69, 92, 100, 151, 165, 179, 182, 198, 218
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte Clone No.: 4115958

<400> 45
ccggaaaggg ggaaaacctt tgggaaattt cccaaaaggg ccttttttaa atccccccaa 60
nangccccnt ttttggggag gaggacctta ancccctttn gaataggggc tcttttaacc 120
cccaagggac ccggccaaag attgtttttt nccaagggat ttgcnccagc catttgggng 180
cnccattgat cccaaccngc atcacatatc cccccggncg gtcattcaac atgcatgact 240
ccagttcggt tgcatctaaa gtgtttagga gttcatatga agacaaaaat ttgttgaaga 300
aaaataagga tgagtcctca gtttccattt ctcacacaaa atgttccttg ttaggagaca 360
tcagtgatgg gaaaaactta atacctaata aatgttttac ttcttttaaa aataatagta 420
aagaaaagtg ttctttgaaa catcaaacaa gaaatcagtg tcagaacaat cctagtgaaa 480
tcatccaaag tacgtatcag gagacacaga acaaaagttc tagtttatca acttcctcaa 540
ttttgtctca gcacaaagaa aataacttag atttgacaag cagattcaag gagcaagaaa 600
tgagcaatgg aattgataaa cagtattcaa attgcaccac tattgacaaa cagatttgta 660
caaataagta taaggaaaaa ataataaatg agaactataa tccaaaattc tttggcaatc 720
ttcagtctga tgattccaaa aaaaatgact caaaaataaa agttactgtg ttggaaatgt 780
ctgaatattt gaacaaatat gaaagcatgt cctcaaataa agactcaaaa aggcctaaga 840
catgtgagca aaatactcaa cttaatagca tagagaatta tctcaataaa gataatgaag 900
gtttcaaatg taaaaagtca gaccaattaa aaaatgaaca agataagcaa gaagatccaa 960
ctaatgaaaa atcccaaaac tattctcaga gaagaagtat aaaagactgt ttgtctacat 1020
gtgagcaacc aaaaaataca gaggtattga ggactacact gaaacattca aatgtgtggc 1080
gaaaacataa ttttcattcc ttggatggaa cttcaaccag agcctttcat cctcaaactg 1140
gattgcctct tctttcaagc cctgttcctc aaagaaaaac acaatcaggt tgctttgatc 1200
tggattcttc attactacat ctgaaaagct tctcatctag aaggaatctg tcttgaacta 1260
tcgtttcgat cctctcggca ttgttgatgg ttttactgcc gaggtagggg caagtggtgc 1320
tttctgcccc acacatttga ctcttccagt tgaagtgtca ttctacagtg tttcagatga 1380
caatgctccc tctccttata tgggtgtgat tactttagag tcccttggta aaaggggtta 1440
tcgagtacct ccttcaggaa caatacaagt ggtatgtgtt ttatagccta caattgcaat 1500
aatcattctc tcacatacat atagaactta gccctttctc ctgctactgc tgctgcaaca 1560
actttgacct tatttaatcc taataagact gtggtgaaga tgtttgttgt gatatatgat 1620
ttacgagata tgccagccaa tcatcagaca ttcctacgac aaagaacttt ttctgtacct 1680
gttaaacaag aagtgaagag aagtgttaat aaagagaaca tccgacacac agaagaacgg 1740
ttattacgct acctcataca tctgaggttc cagagttcta aatctggaaa gatctacctc 1800
catagagacg tacggctcct gttctctaga aagtcaatgg aggttgatag cggtgctgca 1860
tatgaactca aatcttacac tgaatcacca acaaaccctc agttttcacc aagatgttga 1920
taaggagtga tgatttaaag tatttactca gtacccaagt ttgcaagtaa aaattagcat 1980
agaatggagt gtaccaaatt aacaatcagg agagtggatt ctctcctgtt atcctggacc 2040
agtttttatg aaaggattcc tgaaatgaaa tccatatatt ccatgtagac tggaaaaact 2100
catgtcctaa tccttttttgt actgttgaaa ccacttcatt ggacatgttg caatagcaaa 2160
acccccagtt agattagtgt ttacacattt tctcagttat ttaatattta atgttttcct 2220
taatactcaa gtgatgtttg tctctagtgt tctaatgtag cacaaatcct atgtaaaatc 2280
atactatgta tttttgacat taatgttgaa atcaaatata tgcacaagtc tttaattttg 2340
tgtaatgtgt taagtgctgt tcatttaagt tattgaaaat gagaataaaa tgttgagctt 2400
ctttaaaagt aacacactat gcaagcatgt gtacttttta tatctctcat gtttagtttt 2460
tataacacca tatccaggtt gctatctcac atagtagtcc tttaacatat tgtattagca 2520
gtgcaatgtg gactaagctg cttcactttc cctttgcaag ttcagatcat catgcccatt 2580
catagccagg attccttatc cccaaaacag ttctatttttt ccttaatcac tactatagag 2640

tctttacatt aaattactgt cgtatgctag ataatttttct caaattgtta aaagaatatg 2700
tactttggaa acaaattagt atttatattg tga 2733

<210> 46
<211> 2177
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 779255

<400> 46
attgctccaa acgacgccaa cttctactac ctggcccttc acatccgtat acactcaggg 60
gctaagccct agagttgcca cttctgtgaa aaatccttcc gccagctctc ccaccttcag 120
cagcacaccc ggatccactc caagatgcac acggagacca tcaagcccca caagtgcccg 180
cactgctcca agaccttcgc caacacctcc tacctggccc agcacctccg tatccactcg 240
ggggccaagc cctacaactg ttcctactgc cagaaggcct tccgccagct ctcccacctt 300
cagcagcaca cacgaatcca cactggtgat agaccataca aatgtgcaca cccaggctgt 360
gagaaagcct tcacacaact ctccaatctg cagtcccaca gacggcaaca caacaaagat 420
aaacccttca agtgccacaa ctgtcatcgg gcgtacacgg atgcagcctc actagaggtg 480
cacctgtcta cgcacacagt gaagcatgcc aaggtgtaca cctgcactat ctgcagtcgg 540
gcatacacat cagaaacata ccttatgaaa catatgcgca aacacaaccc gcctgatctt 600
cagcaacagg tgcaggcagc agcagcagcg gcagcagtgg cccaggccca ggctcaagct 660
caagcccagg ctcaggctca ggctcaagcc caggcccagg cccaggcctc ccaggcatca 720
cagcagcagc agcagcagca gcagcagcag cagcagcagc agcaacagcc accaccacac 780
ttccagtctc ctggggcagc cccccagggt gggggtggtg gggacagcaa tcccaaccct 840
ccaccccagt gttcctttga cctgaccccg tataagacgg cggagcatca taaggacatc 900
tgcctcactg tcaccaccag caccatccag gtggagcacc tggccagctc ttagagatcc 960
gtgctgccac ccactgggaa gaggaagaag tagtcctggt gtcttctttc tccaactctt 1020
ggtgggaaaa gtccttttct tccttgacag gccttggctc catctccttg ggcctctgtc 1080
acggctttcc ttcacaggat accatccttt ttctgaactc ttcttcaaaa ggaacatcag 1140
ccctcctgat tgcaaaggaa tactgagctg atggtgtcat ccagcagcct cccctcccaa 1200
gcaaagcttc taaaactggg ggtcggtgct caagggaagg atttgctatg acctcataga 1260
accttgtcca gtgtggccac ttaccctatc cttaccctcc ttatcctcaa agtttgggct 1320
gatgtaagac tagaggctgg ccctcccaga taacagagaa aagggagccc caaatgcaac 1380
cagcctcttg ttctattctt gcctgcaaaa gaacagaggt ttctcaaatg cctcagtccc 1440
tgagagccat ttcttccccc acatcgtctc actttgcttc ctattgactg ctggtagaag 1500
gagatttggg gtaggggcta gacctccttt tatttgaagg gggcaagggc tgagatgtgg 1560
tccccaaggg gccagaaatt cccaagttgg tcacaggtgg cttagaagtg tgtgttatgg 1620
ttttacggat ttccttgaag cctctctcct tctctgccta caaagaccct atactctcag 1680
tctccccaac ccacccccaa ggagctgtgg gaggctttgt gttatctgtg aaactccaaa 1740
acaggggtgt tgcggagaag ggagagttca aggcaaacgc aaggactgga cttagctccc 1800
taggtgccac agtcagatgc cggacacgga tttatatata aatatatata tataaatata 1860
ttatacccac tcatcacggc catctttgtt gtaaccattt ctgtgtttat aaatgcatta 1920
tctctgagaa ttttcatatt tgatgttttg tttatttttg tccttttttt ccctctctcc 1980
acccctgtcc tctagccaca gcatttttct ttttgtcttt ttttttttt tttaaatcat 2040
ggcagatttc agaggaaagg aaattaaaaa aaaaatcagg aaaccagttg ttataaagta 2100
atttaaaaat gaagaaaaaa agaaaaaaac ttatgtacaa accaaggggt tttttagaac 2160
attgtataga aatagag 2177

<210> 47
<211> 2685
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 1303605

<400> 47
gctctgtttg acttcatgca gaatttttaa agtttttagg tttcttgaaa atgtaatttc 60
atgaactctt ctaaggctac tgtaactgaa ttccaaccca cagatatgtt acacacgatt 120
ggcatatctt tttcttgcac atcaaggagg ccctcttggc aggccacctc ttgtgttccc 180
cagagcaggc agtggaactc agtgccctcc tggcccagac caagtttgga gactacaacc 240
agaacactgc caagtataac tatgaggagc tctgtgccaa ggagctctcc tctgccacct 300
tgaacagcat tgttgcaaaa cataaggagt tggaggggac cagccaggct tcagctgaat 360
accaagtttt gcagattgtg tcggcaatgg aaaactatgg catagaatgg cattctgtgc 420
gggatagcga agggcagaaa ctgctcattg gggttggacc tgaaggaatc tcaatttgta 480
aagatgactt tagcccaatt aataggatag cttatcctgt ggtgcagatg gccacccagt 540
caggaaagaa tgtatatttg acggtcacca aggaatctgg gaacagcatc gtgctcttgt 600
ttaaaatgat cagcaccagg gcggccagcg ggctctaccg agcgataaca gagacgcacg 660
cattctacag gtgtgacaca gtgaccagcg ccgtgatgat gcagtatagc cgtgacttga 720
agggccactt ggcatctctg tttctgaatg aaaacattaa ccttggcaag aaatatgtct 780
ttgatattaa aagaacatca aaggaggtgt atgaccatgc caggagggct ctgtacaatg 840
ctggcgttgt ggacctcgtt tcaagaagca accagagccc ttcacactcg cctctgaagt 900
cctcagaaag cagcatgaac tgcagcagct gcgagggcct cagctgccag cagacccggg 960
tgctgcagga gaagctacgc aagctgaagg aagccatgct gtgcatggtg tgctgcgagg 1020
aggagatcaa ctccaccttc tgtccctgtg gccacactgt gtgctgtgag agctgcgccg 1080
cccagctaca gtcatgtccc gtctgcaggt cgcgtgtgga gcatgtccag cacgtctatc 1140
tgccaacgca caccagtctt ctcaatctga ctgtaatcta atctgttgtg cttttgttgg 1200
acttggcatg tttccatgaa ctgcactatt ataaactatt aaaatgatag attgtggaga 1260
aagtaattat tccaacaccc atctgccatg cgatgttaaa aaaaaaaaaa aggaagaaaa 1320
ataacacagc tactcctcac tgcaaaaaca tatccatgcg tagaatcaac aactccagtc 1380
atgggaccag gaggagctct gggacgcaga cacattcctt ggatgttgat ttttttatg 1440
atctagtaaa ggaataggta aagtctttga tgtcagtgaa gtggcaacat agccaaaaag 1500
ttgggtacct tttaggaaat gatgttgtaa gtctccttaa tgtatcctga ggtaagtttc 1560
ctactggcag cagattttgt aagaattact tttaagaatt tcattctttt tgtatggtca 1620
tggagctcca accatttta ataggaaagt cttttgtaaa ttgttgtcgt tttaatgtca 1680
tttctgtctt tataacttga tcaagaatga ttggaaggca aacaggttta caaatcaatt 1740
ctgtgacttt taaaaagttg acaatgttgt cagatttaaa ccagtgtggc tagtaaaaag 1800
cagctcactc aatgtgggtg gctccctatt cctttacgct ccccctatcc ctaccccaca 1860
agcctttcga ttataaaata ctaccaatct tgttataaga ttactgtgga gtagtcaagt 1920
actccccggg ccttctgagc tggtggaata ttttatttca gactgaaaac agagagcact 1980
ctccttggga agggaaagcg gagcttgctg agtgagagat ggagcctcat ggtgtacaac 2040
tgagggtagt taactcatca cttctcccaa gcactcgatc ccagcttcac ccactggtgt 2100
tgctttgctt gaactgttca agccttttat agccttacca taagtattta gatatggtgt 2160
ccttttctgt ttttgggggg ggagttttgt tgtgtttttt taaagtaagt gcttaagtat 2220
taactttggg ttgtcccctc tgtatgtttc gaaggggttt tggttctttt tgcttctgtt 2280
ttcttaaaca tgttttccac tcccacttgg gcattttgga agctggtcag ctagcaggtt 2340
ttctgggatg tcgggagacc tagatgacct tatcgggtgc aatactagct aaggtaaagc 2400
tagaaaccta cactgtcact ttactgagat ttctgagtat acttttcata ttgccttaat 2460
gtagcagtaa tgtgtttatg catttgtttc tttgcacaga cattttgtca aatattaaaa 2520
ctctactttt ttatggcaca tattagcata taagccttta ttccaagagg tatttatttt 2580
ttcacttgta aaaaaataat gtttccacgt aaagaactct gttatatcct agaggactct 2640
gtcttttata ttcgggataa taaagacttt aaagcaaaaa aaaaa 2685

<210> 48
<211> 2408
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 1611167

<400> 48
ggctgcatcc cggccggcgg ttccggagcc tcgcgtctgg ggaggcgcgg gccggcacgc 60
tcgagcttag ggagatctgc cttctggaga ctgcgccgtc ctcccgggag agccagaaag 120
aggacatggc tgctgggcag cgggaagcga ggccccaggt gtcactgacg tttgaggatg 180
tggctgtgct gtttacccga gatgagtgga gaaagctggc cccttctcag agaaacttgt 240
accgggatgt gatgctggag aactatagga acctggtctc actggggctc ccatttacca 300
aaccaaaagt gatctccctg ttgcagcaag gagaagatcc ctgggaggtg gagaaagacg 360
gttctggcgt ctcctctcta ggatcgaaga gcagtcataa aaccacaaag tcaacgcaaa 420
cacaagactc ttcatttcag ggactgatac tgaaaagatc caacaggaat gtaccttggg 480
atttgaaatt agaaaagcct tacatatatg aaggcagatt agagaaaaag caggataaaa 540
agggaagttt tcagatagtt tcagccaccc acaaaaaaat ccccactata gaaagaagcc 600
ataaaaatac tgaattgagc caaaacttca gcccaaagtc agtgcttatt aggcaacaga 660
tacttcccag agaaaaaaca ccaccaaaat gtgaaataca aggaaacagc ctcaaacaga 720
attcacaatt acttaatcaa ccaaaaatta cagcagataa acgctataaa tgtagtctgt 780
gtgaaaaaac cttcattaac acttcatccc ttcgtaaaca tgagaaaaac catagtggag 840
agaaactatt taagtgtaaa gaatgttcaa aagcctttag ccaaagttca gctcttattc 900
aacatcaaat aacacatact ggagagaaac cctacatatg taaagaatgt gggaaagcct 960
ttactctcag tacatccctt tataaacatc taagaaccca tactgtggag aaatcctaca 1020
gatgtaaaga atgtggtaaa tccttcagcc gaaggtcagg ccttttttata catcaaaaaa 1080
ttcatgctga agaaaaccct tgtaagtata atccgggtag gaaggcatct agttgcagca 1140
catccctttc tggatgtcaa agaattcatt ctagaaagaa gtcctactta tgtaatgaat 1200
gtggcaacac ctttaagtct agctcatccc ttcgttatca tcagagaatt cacactggag 1260
agaagccttt taaatgtagt gaatgtggga gagccttcag ccagagtgcc tctcttattc 1320
aacatgaaag aattcacacc ggagaaaagc cctatagatg caatgaatgt gggaaaggct 1380
ttacttctat ttcacgactt aatagacacc gaatcattca tactggagag aagttttata 1440
attgtaatga atgtggtaaa gccttaagct cccactcaac acttattatt cacgagcgaa 1500
ttcatactgg agaaaaacca tgtaaatgta aagtatgtgg aaaagccttc agacagagtt 1560
cagctctcat tcaacatcag agaatgcata ctggagaaag accctataaa tgtaacgagt 1620
gtgggaaaac attcaggtgt aactcatcac ttagtaatca ccagagaatt catactggag 1680
agaaaccata tcgatgtgag gaatgtggga tatcttttgg ccaaagttca gctcttattc 1740
agcatcgaag gattcataca ggagaaaaac cctttaaatg taatacatgt ggaaaaactt 1800
ttagacaaag ctcatcacgt attgcacatc agagaattca tactggagag aaaccctatg 1860
aatgtaatac atgtgggaaa cttttcaacc ataggtcatc ccttactaat cattataaaa 1920
ttcatatcga agaggacccc tagaaagtag atttgtatgt gtgaaagcct taaaccaaag 1980
ctcatcgaag aatacatcct tgagagagat gtaataaatg taatggatgt gaaaaaaact 2040
gtaataattt agccctcatt aggtatttaa ttccatggat aaacctcagc tatataatag 2100
ataatgagga aagtgtttgt gcctgtcaga cacttaaaaa aataacctga gatgaagaat 2160
ttacaattga agacattgac tttagccatt tgtgaaatgg gtttgctttt tccctttcc 2220
tacagacgta tatgctagat gtcacgtgat catcagaaac agatatccga gtgggtgggg 2280
aggtgtgctt tagatttctc attagaagac caccaaactg gtaatatttt tatagcattt 2340
taatagcata caaatgaatt gatcaaattg taccttttta gagaaaagga ccaaaataaa 2400
agaaaaat 2408

<210> 49
<211> 2990
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 1907472

<400> 49
gtacaacttt tgcaattcca gttgtgtggc taaatttcag gctctaagta tgcagtcatc 60
tccaaatggc cagtttgtag cgccaagtga tattcagttg aaatgcaact actgcaaaaa 120
ttccttttgt tcaaaaccag aaatcctgga atgggagaac aaagtgcatc agttctgcag 180
caaaacttgt tcagatgact ataagaagtt gcattgcata gttacatatt gcgaatactg 240
tcaagaggag aagactcttc atgaaacagt aaatttctct ggcgttaaga gacctttctg 300
tagtgaaggc tgcaaattat tatacaaaca ggattttgcc agacgtttag gattgagatg 360
tgttacttgc aactattgtt ctcagctatg taagaaggga gcaactaaag aactcgatgg 420
tgttgtgaga gatttctgca gtgaagattg ctgtaaaaaa tttcaggatt ggtactacaa 480
ggctgcaagg tgtgactgtt gtaaatctca aggaactctt aaagagcgag ttcagtggcg 540
tggggaaatg aaacatttct gtgatcaaca ttgcttactg cgtttctact gtcaacaaaa 600
tgagcccaac atgacaactc agaaaggacc tgaaaactta cattatgatc agggttgtca 660
gacatctcga accaaaatga caggttcagc accacccccc tctccaacac ctaacaaaga 720
gatgaagaac aaagcagttc tttgcaaacc tttaacaatg acaaaagcta cttactgtaa 780
acctcacatg cagaccaaat cttgtcagac agatgatact tggaggacag aatatgttcc 840
agtgcctatc cctgtgcctg tgtatatccc agttcctatg cacatgtaca gtcagaatat 900
tcctgttcct actacagttc ctgttcctgt gccagttcct gtttttctgc ctgctccatt 960
ggacagcagt gagaagattc ctgcagcaat tgaggagcta aaaagcaagg tttcttcaga 1020
tgctcttgat acagagttgc ttacaatgac ggatatgatg agtgaagacg aggggaaaac 1080
agagacaacc aacatcaaca gtgtaattat tgaaacagat ataattggtt cagacctttt 1140
gaagaactct gacccagaga cacagtccag catgcctgat gtaccatatg aaccagattt 1200
ggatatcgaa atagattttc ccagagctgc tgaggagctt gatatggaaa atgaattttt 1260
attaccacct gtttttggcg aagaatatga ggaacagccc agacctcgat ctaaaaaaaa 1320
gggagccaag agaaaggctg tatcaggata ccagtctcat gatgatagtt ctgacaattc 1380
agaatgcagc tttcctttca aatatacgta tggcgtaaat gcatggaaac actgggtcaa 1440
aactaggcaa cttgatgaag atcttctggt attagatgag ttaaaatctt ctaaatcagt 1500
aaagttaaaa gaggatctac tctctcacac cacagctgag cttaactatg ggttagctca 1560
ttttgtcaat gagatccgac ggccaaatgg agagaattat gcacctgaca gcatctatta 1620
cctttgcctt ggaatacagg agtatttgtg tggaagtaat cgaaaagaca acatatttat 1680
tgatcctgga taccaaacat ttgagcaaga attgaataaa atactgcgaa gctggcaacc 1740
aagcatactt ccagatgggt caatattctc tcgagttgaa gaagactatc tctggaggat 1800
aaaacaacta ggatcacact ctccagtagc tcttctgaat acactgttct actttaacac 1860
taagtatttt ggcctgaaaa cagtggaaca acacttaaga ctttcctttg gcactgtgtt 1920
taggcattgg aaaaaaaatc ctttaacgat ggaaaacaaa gcgtgtcttc gataccaagt 1980
gtcttccttg tgtggaacag ataatgaaga taaaattact actggaaaaa gaaaacatga 2040
agatgatgag ccagtatttg aacaaattga aaacacagcc aatccttcca gatgtcctgt 2100
gaaaatgttt gaatgctact tgtctaaaag tccacagaat cttaatcaga ggatggatgt 2160
tttttatttg caaccagaat gctctagttc tacagatagc cctgtctggt atacgtctac 2220
ttcactggac cgaaacacct tggaaaatat gcttgtacgg gttcttctag taaaagatat 2280
ttatgataaa gacaattatg aactggatga agacacagac taaaaaggaa cgttgcagaa 2340
gcaatcggga taaaacagca ttagatagtc atgctgctag atctttatta tggaaaacat 2400
ttcaagttta ctccttctgt tttgagtttt gtagcagtgt acccacgctg ggtattacca 2460
tgtaaataat ctgtgagtga aagttgccat tattctatgt agtggtttta ggatacttaa 2520
caaatacatt caaattcttt ttttattatt atttatttga ttaggtatgt ttgtaacttt 2580
ttacattaca gaatatgaat gagaatgtgc catgtataat ttttttcttg tagtaagaaa 2640
catccatatt gcacaactct actgttgcaa agcttccttg gaaggggggct cttttactgg 2700
gttcttaacc agatggttgt gtatgggtag cactactaaa agtttagaac ttgcagtgtc 2760
tttcggaatt tttaaaataa actgtaaact aataggctgg ggtttttgtt ttgttttggg 2820
gttttgtttt gtttggtttt acatttttagt tactgaagcc ttacaaggtt atgtagagag 2880
ataccatctt ctgtaccaaa aatagacaag agaatgctgt caatattggt gtactgtaat 2940
gtgaatctat gctggtgaaa acaatttttt ttccccttat taaaaccttа 2990

<210> 50
<211> 771
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 1985458

<400> 50
ccctactaa agccttgggg ttagtacgcg tgcgcacagt ttcttccgac agttgtgttg 60
tgccaatggt ggagaagaaa acttcggttc gctcccagga ccccgggcag cggcgggtgc 120
tggaccgggc tgcccggcag cgtcgcatca accggcagct ggaggccctg gagaatgaca 180
acttccagga tgacccccac gcgggactcc ctcagctcgg caagagactg cctcagtttg 240
atgacgatgc ggacactgga aagaaaaaga agaaaacccg aggtgatcat tttaaacttc 300
gcttccgaaa aaactttcag gccctgttgg aggagcagaa cttgagtgtg gccgagggcc 360
ctaactacct gacggcctgt gcgggacccc catcgcggcc ccagcgcccc ttctgtgctg 420
tctgtggctt cccatccccc tacacctgtg tcagctgcgg tgcccggtac tgcactgtgc 480
gctgtctggg gacccaccag gagaccaggt gtctgaagtg gactgtgtga gcctgggcat 540
tcccagagag gaagggccgc tgtgcactgc ccggccttca gaaagacaga atttcatcac 600
ccaatgcagg gggagctctt cctggaccaa gggaggagcc gctcattcac ccaacaaaac 660
tgtgtcttat ctgccaggaa agaccagcct cactcctggg aactgtctgg caggtaggct 720
gggccccccа gtgctgttag aataaaaagc ctcgtgccgg aaaaaaaaaa a 771

<210> 51
<211> 2076
<212> DNA
<213> Homo sapiens

<220>
<221> unsure
<222> 1957
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte Clone No.: 2726431

<400> 51
cccgcctgcc cctagcggtc cctggcgtcc cggttcctgt cgcgctcacc tgcgccggta 60
gcgaagaaat cgccccggga catggactca gtggtctttg aggatgtggc tgtggacttc 120
accctggagg agtgggcttt gctggattct gctcagaggg acctctacag agatgtgatg 180
ctggagacct ttcagaacct ggcctcagta ggaaaaattt gggacagtct tagcatcgaa 240
gatcaaacca caaaccaggg gagaaatctc agtagaaatc atgggttgga gagactctgt 300
gaaagtaatg atcaatgtgg agaagccctc agccagattc cacatcttaa tctgtacaag 360
aaaattccac ctggagtaaa acagtatgaa tacaacacgt acggaaaagt cttcatgcat 420
cgccgcacat ccctcaagag tcccatcaca gttcacactg gacacaaacc atatcagtgc 480
caggaatgtg ggcaggccta cagttgtcgt tcacacctaa gaatgcatgt gagaacccac 540
aatggagaga gaccctatgt gtgtaaatta tgtgggaaaa cctttcctcg tacttcctcc 600
ctcaatcggc atgtaaggat tcacactgct gagaaaactt acgaatgtaa gcaatgtggg 660
aaagccttta ttgacttctc aagtcttact agtcatctca gaagtcacac cggagagaag 720
ccatataagt gtaaggaatg tgggaaagct ttcagttatt cctcaacgtt tcgaagacac 780
acaataacac acactggcga gaagccatat aaatgtaagg aatgtgcgga agcctttagt 840
tattcctcaa cttttcgaag acatatgatt tcacacactg gagagaagcc acataaatgt 900
aaagaatgtg gggaggcctt cagttattct tcggcttttc gaagacacat gataacacac 960
actggagaga aaccctacga atgcaaacaa tgtgggaaaa ccttcattta tctccagtcc 1020
tttcgaagac atgaaaggat tcacactgga gagaaaccct acgaatgcaa acagtgtggg 1080
aagaccttca tttatcccca gtcctttcga agacatgaaa ggactcatgg tggagagaaa 1140
ccctatgaat gcaaccagtg cgggaaagca ttcagtcacc cctcctcctt tcgaggacac 1200
atgagggtgc acactggaga gaaaccctat gagtgcaagc aatgtgggaa aactttcaat 1260
tggcccatat ctttacgaaa acatatgaga acacatacta gagagaaacc ctatgaatgt 1320

aagcagtgtg ggaaagcctt cagcttgtct gcttgctttc gagaacatgt gagaatgcac 1380
cctgaagaca aatcctatga atgcaagcta tgtgggaaag ctttctattg ccacatatcc 1440
ttacaaaaac atatgagaag gcataccgca gagaaactct ataaatgcaa gcagtgtggg 1500
aaagctttca gttggcctga acttttgcaa caacatgtga gaacgcacac tgtagagaag 1560
ccctatgaat gtaaggaatg tgggaaggtc ttcaaatggc catcatcttt accaatacat 1620
atgagactgc acactggaga gaaaccttat caatgtaagc attgtgggaa agcattcaat 1680
tgttcctcat ccttaaggcg acatgtgaga atacacacta cagaaaaaca gtataagtgt 1740
aatgtaggac atcctcctgc aaatgaattc atgtgcagtg cttcagaaaa gtcacaccag 1800
gagagagatc tgatcaaagt tgtaaatatg gtgttgcctt tatgagttcc ttatcctgaa 1860
agtggacact caaggagtgt gtctgtagtt catttgcata gaaactatag cgaagaggcc 1920
cgcaccgatc gcccttccca acagttgcgc agctganatg gcgaatggga cgcgccctgt 1980
agcggcgcat taagcgcggc gggtgtggtg gttacgcgca gcgaccgcta cacttgccag 2040
cgccctagcg cccgctcctt tcgctttctt cccttc 2076

<210> 52
<211> 1197
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 2743828

<400> 52
cccgggtccc cttctccagg caggaagatg tccaagcccc gcgcggtgga ggcggcggcg 60
gcggcggcgg cggtggcagc gacggccccg ggcccggaga tggtggagcg gaggggcccg 120
gggaggcccc gcaccgacgg ggagaacgta tttaccgggc agtcaaagat ctattcctac 180
atgagcccga acaaatgctc tggaatgcgt ttcccccttc aggaagagaa ctcagttaca 240
catcacgaag tcaaatgcca ggggaaacca ttagccggaa tctacaggaa acgagaagag 300
aaaagaaatg ctgggaacgc agtacggagc gccatgaagt ccgaggaaca gaagatcaaa 360
gacgccagga aaggtcccct ggtacctttt ccaaaccaaa aatctgaagc agcagaacct 420
ccaaaaactc caccctcatc ttgtgattcc accaatgcag ccatcgccaa gcaagccctg 480
aaaaagccca tcaagggcaa acaggccccc cgaaaaaaag ctcaaggaaa aacgcaacag 540
aatcgcaaac ttacggattt ctaccctgtc cgaaggagct ccaggaagag caaagccgag 600
ctgcagtctg aagaaaggaa aagaatagat gaattgattg aaagtgggaa ggaagaagga 660
atgaagattg acctcatcga tggcaaaggc aggggtgtga ttgccaccaa gcagttctcc 720
cggggtgact ttgtggtgga ataccacggg gacctcatcg agatcaccga cgccaagaaa 780
cgggaggctc tgtacgcaca ggacccttcc acgggctgct acatgtacta ttttcagtat 840
ctgagcaaaa cctactgcgt ggatgcaact agagagacaa atcgcctagg aagactgatc 900
aatcacagca aatgtgggaa ctgccaaacc aaactgcacg acatcgacgg cgtacctcac 960
ctcatcctca tcgcctcccg agacatcgcg gctggggagg agctcctgta tgactatggg 1020
gaccgcagca aggcttccat tgaagcccac ccgtggctga agcattaacc ggtgggcccc 1080
gtgccctccc cgccccactt tcccttcttc aaaggacaaa gtgccctcaa agggaattga 1140
attttttttt tacacactta atcttagcgg attacttcag atgtttttaa aaagtat 1197

<210> 53
<211> 2843
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 2998209

<400> 53

tcagatgtac ccagtgctta tcccaacgta ggaccaggag tcctacagaa tctcctcagc 60
ctcaggatct ggctcaagac cagttgcttt ttctgttgat ccttgtgagg aatacaatga 120
tagaaataaa ctgaacacta ggacacagca ctggacttgc tctgtttgca catatgaaaa 180
ctgggccaag gctaaaagat gtgttgtttg tgatcatccc agacctaata acattgaagc 240
aatagaattg gcagagactg aagaggcttc ttcaataata aatgagcaag acagagctcg 300
atggagggga agttgcagta gtggtaatag ccaaaggaga tcacctcctg ctacgaagcg 360
ggactctgaa gtgaaaatgg attttcagag gattgaattg gctggtgctg tgggaagcaa 420
ggaggaactt gaagtagact ttaaaaaact aaagcaaatt aaaaacagga tgaaaaagac 480
tgattggctc ttcctcaatg cttgtgtggg ggttgtagaa ggtgatttag ctgccataga 540
agcatacaag tcatcaggag gagacattgc acgtcagctc accgcagatg aagtacgctt 600
gctgaatcgt ccttctgcct ttgatgttgg ctatactctt gtacacttgg ctatacgttt 660
tcagaggcag gatatgctag caatattgct tacagaggtg tctcaacaag cagcaaagtg 720
tattccagca atggtgtgtc ctgaactgac agaacaaatc cggagagaga tagctgcctc 780
tcttcatcag agaaaggggg attttgcttg ctattttctg actgaccttg taacatttac 840
attgccagca gatattgaag atttgccccc aacagtccaa gaaaaattat ttgatgaggt 900
gcttgataga gacgttcaaa aagaattaga agaagaatct ccaattatta actggtcctt 960
ggaattggct acacgtttgg acagtcgact gtatgcactt tggaaccgga ctgcaggaga 1020
ctgcctactt gattcagttc tacaagctac ctggggcatc tatgacaagg actcagtgct 1080
tcggaaagcc ctgcatgaca gcctgcatga ctgttcacat tggttttaca cacgctggaa 1140
agattgggaa tcatggtatt ctcagagctt tggtttacat ttttccttga gagaagaaca 1200
gtggcaagaa gactgggcat ttatactctc tcttgctagt cagcctggag caagcttgga 1260
gcagacgcac atttttgtac tggcacatat tcttagacga ccaattatag tttatggagt 1320
aaaatattac aagagtttcc ggggagaaac tttaggatat actcggtttc aaggtgttta 1380
tctgcctttg ttgtgggaac agagtttttg ttggaaaagt ccgattgctc tgggttatac 1440
gaggggccac ttctctgctt tggttgccat ggaaaatgat ggctatggca accgaggtgc 1500
tggtgctaat ctcaataccg atgatgatgt caccatcaca tttttgcctc tggttgacag 1560
tgaaaggaag ctactccatg tgcacttcct ttctgctcag gagctaggta atgaggaaca 1620
gcaagaaaaa ctgctcaggg agtggctgga ctgctgtgtg acggaggggg gagttctggt 1680
tgccatgcag aagagttctc ggcggcgaaa tcacccccctg gtcactcaga tggtagaaaa 1740
atggcttgac cgctaccgac agatccggcc gtgtacatcc ctgtctgatg gagaggaaga 1800
tgaggatgat gaagatgaat gaaaaaaaaa atcaaacagc agaagaccaa ggcatcagat 1860
ctgtaatgac cctaaagtta gtgtggtgct ccaagcagag tcgacatcat ggaatgaacc 1920
aaatctggca ggatctgctc ggggaagtgt tttcctggac cacacacacc ttatggagat 1980
aatgcctctg ctgcgtgagg agacagagaa ctttagttgg actacagttt gtaaaaaaaa 2040
ctaattttat taagacagaa ctttttttcc ttccaaattg taaatctgtc tataaatgta 2100
acgcatgtgg ttgtgtaaga cattgtttaa taggaaaagt tgtaccagca tcttcatatt 2160
attgagaaaa ttttttccag catgggcact tagaaaaagc acatggcaaa tggctctttg 2220
ttcctttcag atattatttc agtagaacct ggcattctcc tttcacctta aaagatccat 2280
ctaagtctca gatctggaaa cgttttgtac cgattatcca cagcaaaaca aaaataagct 2340
tttattttat taataatttc gttcctcttg tgcccaatca aatcttttag gaacaaactg 2400
caagaaaagc taagaatgtt ttagagtgaa ctaaatacag acattgctta cttgttttga 2460
agagggtttt ggttttggtt attgtgtctt ttaagttttc tgatatgccc cctttcaata 2520
tttagatatt tatttgttgg gaagaatacc ttaaaatgag ggttcttatt ccagattctg 2580
ggcagtggtc tgtgagtagt ttttttcctg gatgaaaagg gagcaagccc acttgtcact 2640
aaatgaattg tgtgaaatgt gctcacttgg actccatcaa caatgtgctg ctcccagatt 2700
gccatgccag agggtcttcg gattcttcct tctatcacct ctgctctaag caaatcttgt 2760
tagaagggca tgcctttgct taggcagatt gggaatacca attcactaca gaataaagat 2820
tttaaaaatg caaaaaaaaa aaa 2843

<210> 54
<211> 1272
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature

<223> Incyte Clone No.: 3340296

<400> 54
tctacggcca cgactctggg agtggggaaa cagagagccg gttcctctgc tgcagaagtc 60
ctcggggttc cttctcacaa ctctgcgaag gggaaagggt tgtgagaccc aaccagaccc 120
caactccagc tcccagcagg aggtggctgc gccacactcg ggaggcctct tggtttcagg 180
gtctctctgt ctctctctca ccctcttcct cgctttctct gtctctctgt ctctctctct 240
ctctctccct cgtccactcc cccaaacatg tccaccggct ccctcagcga tgtggaggac 300
cttcaagagg tggagatgtt ggaatgtgac gggttgaaaa tggattcgaa caaggaattt 360
gtgacttcca acgagagcac cgaggagagc tccaactgcg agaatgggtc tccccagaag 420
ggccgcggcg gcctgggcaa gaggaggaag gcgcccacca agaagagccc cctgagcggg 480
gtcagccagg aggggaagca ggtccagcgc aacgccgcca acgcgcgaga gcgggcccgc 540
atgcgagtgc tgagcaaggc cttctccaga ctcaagacca ccctgccctg ggtgcccccc 600
gacaccaagc tctccaagct ggacacgctc aggctggcgt ccagctacat cgcccacttg 660
aggcagatcc tggctaacga caaatacgag aacgggtaca ttcacccggt caacctgacg 720
tggcccttta tggtggccgg gaaacccgag agtgacctga aagaagtggt gaccgcgagc 780
cgcttatgtg gaaccaccgc gtcctgacct tggaggtgcg agtctgggaa aggcgcgctc 840
ccgggggggag cgggccccgg gaaggcgacc cctgccctca gtgctctctg tctctgcttc 900
cccctcgcaa tgctcctctc tctgtcccac cccgcgagaa cactttacaa cgacgaggag 960
attcgtttcc aaaccagagg agatcaattg tacttacaaa gattcccatc tatttaactt 1020
tattaacttc taccgtgaat gactctgcaa gccttgctgg tccaagtgca atatgtaatt 1080
ataaatatat aaatagataa gagcctatca atgtatcttt tgtacaatat gttgtaaaat 1140
gtagatcata ggatagctga ctttgacagt cacatttata aagtaattca cttaaagata 1200
tatatttttt tcaaacaagt tttgctactt ttgaaaataa atctttcttt atattgctaa 1260
aaaaaaaaaa aa 1272

<210> 55
<211> 1117
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 3536740

<400> 55
tgcactgcca cgccgagctg aggctgagct cgcccggcca gctcaaagca gccaggcggc 60
gctacaagac tttcatgatc gacgagatcc tctccaagga gacctgcgat tactttgaga 120
aactttccct ctactccgtg tgcccgtcgc tggtcgtgcg acccaagccc ctgcattcct 180
gtacgggctc cccttccctg cgggcatatc cgctcctctc ggtgatcacc cgccagccca 240
ctgtcatctc ccacctggtc cctgccaccc cgggaatcgc ccaggcactg tcctgccacc 300
aggtcaccga ggcggtctct gctgaggccc caggggggcga ggccctagcc agcagcgagt 360
cagagacgga acagcccacg ccccgacaga agaagccccg ccggagtcgc accatcttca 420
ccgagctgca gctcatgggc ctggagaaga aattccagaa gcagaagtat ttgtcaaccc 480
cagacaggtt ggacttggct cagtctctgg gactcactca gctgcaggtg aagacctggt 540
atcagaatcg caggatgaaa tggaagaaaa tggttcttaa aggtggacag gaagcaccca 600
caaaacccaa aggtcgcccc aagaagaact ccatccccac atcagaagag attgaagctg 660
aagagaagat gaacagccag gcccagggtc aggagcagct ggagccctct caggggcagg 720
aggagctctg tgaagcacag gaaccgaaag cacgtgatgt ccccttagag atggcagagc 780
caccagaccc gccccaggag ttgccaatac cctcttcgga acccccacca ttaagctaaa 840
gtaaaaccct tttgagggaa gagggagact ggggagaagg gaaaagagag aaggcaggga 900
gagtagggag agaaaaacctt ccagcagccc agtaaactgc gggcgaagag atctacccgt 960
ctccctccct cccacagtta ccattggcct tgtcatcgca agcatttgac aaagacttgc 1020
ttgtcttggg cctgtcacct cctgaaaggc tgctttagct gtggatgccc ttgattaagg 1080
gagagagcgc ctaggagctg cctgccccag ctggggt 1117

<210> 56
<211> 3033
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 082155

<400> 56
gcagatgcta gaacataatg tagcattact ttccccaggg tttattgtta tgtaagttct 60
tgttcagctt cctttgtttt ctttcacttc tgagaattta actttcgttt ctcactcagc 120
tcctgtgggg aaactcattt gtggagacca gccctctggc ttggtgagtg aatctggttt 180
acaccggctc ctgccctgcc ttcactcttc tcccctgatt caagactcct ctgctttgga 240
ctgaagcact gcaggagttt gtgaccaaga acttcaagag tcaagacaga aggaagccaa 300
gggagcagtg caatggattt ctcagtaaag gtagacatag agaaggaggt gacctgcccc 360
atctgcctgg agctcctgac agaacctctg agcctagatt gtggccacag cttctgccaa 420
gcctgcatca ctgcaaagat caaggagtca gtgatcatct caagagggga aagcagctgt 480
cctgtgtgtc agaccagatt ccagcctggg aacctccgac ctaatcggca tctggccaac 540
atagttgaga gagtcaaaga ggtcaagatg agcccacagg aggggcagaa gagagatgtc 600
tgtgagcacc atggaaaaaa actccagatc ttctgtaagg aggatggaaa agtcatttgc 660
tgggtttgtg aactgtctca ggaacaccaa ggtcaccaaa cattccgcat aaacgaggtg 720
gtcaaggaat gtcaggaaaa gctgcaggta gccctgcaga ggctgataaa ggaggatcaa 780
gaggctgaga agctggaaga tgacatcaga caagagagaa ccgcctggaa gaattatatc 840
cagatcgaga gacagaagat tctgaaaggg ttcaatgaaa tgagagtcat cttggacaat 900
gaggagcaga gagagctgca aaagctggag gaaggtgagg tgaatgtgct ggacaacctg 960
gcagcagcta cagaccagct ggtccagcag aggcaggatg ccagcacgct catctcagat 1020
ctccagcgga ggttgacggg atcgtcagta gagatgctgc aggatgtgat tgacgtcatg 1080
aaaaggagtg aaagctggac attgaagaag ccaaaatctg tttccaagaa actaaagagt 1140
gtattccgag taccagatct gagtgggatg ctgcaagttc ttaaagagct gacagatgtc 1200
cagtactact gggtggacgt gatgctgaat ccaggcagtg ccacttcgaa tgttgctatt 1260
tctgtggatc agagacaagt gaaaactgta cgcacctgca catttaagaa ttcaaatcca 1320
tgtgattttt ctgcttttgg tgtcttcggc tgccaatatt tctcttcggg gaaatattac 1380
tgggaagtag atgtgtctgg aaagattgcc tggatcctgg gcgtacacag taaaataagt 1440
agtctgaata aaaggaagag ctctgggttt gcttttgatc caagtgtaaa ttattcaaaa 1500
gtttactcca gatatagacc tcaatatggc tactgggtta taggattaca gaatacatgt 1560
gaatataatg cttttgagga ctcctcctct tctgatccca aggttttgac tctctttatg 1620
gctgtgcctc cctgtcgtat tggggttttc ctagactatg aggcaggcat tgtctcattt 1680
ttcaatgtca caaaccacgg agcactcatc tacaagttct ctggatgtcg cttttctcga 1740
cctgcttatc cgtatttcaa tccttggaac tgcctagtcc ccatgactgt gtgcccaccg 1800
agctcctgag tgttctcatt cctttaccca cttctgcata gtagcccttg tgctgagact 1860
cagattctgc acctgagttc atctctactg agaccatctc ttcctttctt tccccttctt 1920
ttacttagaa tgtctttgta ttcatttgct agggcttcca tagcaaagca tcatagattg 1980
ctgatttaaa ctgtaattgt attgccgtac tgtgggctgg aaatcccaaa tctagattcc 2040
agcagagttg gttctttctg aggtctgcaa ggaagggctc tgttccatgc ctctctcctt 2100
ggcttgtaga aggcatcttg tccctatgac tcttcacatt gtctttatgt acatctctgt 2160
gcccaagttt tccctttta ttaagacacc agtcatactg gctcagggcc caccgctaat 2220
gccttaatga aatcatttta acattatatt ctctacaaag accttatttc caaataagat 2280
aatatttgga ggtattggga ataaaaactc caacatataa atttgaggaa ggcacgattt 2340
cactcataac aatcttaccc tttcttgcaa gagatgcttg tacattattt tcctaatacc 2400
ttggtttcac tagtagtaaa cattattatt ttttttatat ttgcaaagga aacatatcta 2460
atccttccta tagaaagaac agtattgctg taattccttt tcttttcttc ctcatttcct 2520
ctgccccttа aaagattgaa gaaagagaaa cttgtcaact catatccacg ttatctagca 2580
aagtacataa gaatctatca ctaagtaatg tatccttcag aatgtgttgg tttaccagtg 2640
acaccccata ttcatcacaa aattaaagca agaagtccat agtaatttat ttgctaatag 2700
tggattttta atgctcagag tttctgaggt caaattttat cttttcactt acaagctcta 2760
tgatcttaaa taatttactt aatgtatttt ggtgtatttt cctcaaatta atattggtgt 2820

tcaagactat atctaattcc tctgatcact ttgagaaaca aacttttatt aaatgtaagg 2880
cacttttcta tgaattttaa atataaaaat aaatattgtt ctgattatta ctgaaaagat 2940
gtcagccatt tcaatgtctt gggaaacaat tttttgtttt tgttctgttt tctttttgct 3000
tcaataaaac aatagctggc tctaaaaaaa aaa 3033

<210> 57
<211> 6138
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 095477

<400> 57
ggctgaggac tgactggggt tctgagactc cctgtcccgg accgcagcgt taaaaggatc 60
tgaacaaagt ctgctcaaat ctcctgctgt gaaccagcag aatttttgaa caggtttctt 120
cacatataaa aatctattgt aaaaatacgg aaaagaatgg cagcggaaac gcagacactg 180
aactttgggc ctgaatggct ccgagctctg tccagtggtg ggagtattac atcccctcct 240
ctttctccag cattgccgaa gtataaatta gcagattatc gttacggcag agaagaaatg 300
ttagcacttt tccttaaaga caacaagata ccttcagacc ttctggataa agaatttctg 360
cctatcctcc aggaggaacc ccttccacca ttggctctgg tacccttac agaagaagaa 420
cagagaaact tttccatgtc tgtaaatagt gctgctgtcc tgcgattgac aggacgagga 480
ggaggaggaa cagtggtggg ggctcctaga ggtcgaagtt cttcaagagg gcgaggcaga 540
ggcagaggtg aatgtggttt ctaccaaaga agttttgatg aagtagaggg tgtttttggt 600
cgaggaggtg gcagagaaat gcatagatcg cagagctggg aggaaagggg tgacagacgt 660
tttgaaaaac caggacgaaa agatgtaggg agaccaaatt ttgaggaagg tggaccaaca 720
tcagtaggga gaaagcatga atttatacgc tcagaaagtg aaaattggcg catctttaga 780
gaggaacaaa atggagaaga tgaagatgga ggttggcgac tagctggatc aaggagggat 840
ggagagaggt ggcgacctca cagtcctgat ggccctcgtt ctgcaggctg gcgggaacac 900
atggaacgac gtcggaggtt tgagtttgat tttcgagata gagatgatga acggggttac 960
cgaagggttc gctctggcag tgggagcata gatgatgaca gggatagctt gcccgaatgg 1020
tgcttagagg atgctgaaga agaaatgggt acatttgact catctggagc attccctttct 1080
ctaaaaaaag tacagaaaga gcctattcca gaagagcagg agatggactt ccggcctgtg 1140
gacgaagggg aggagtgctc tgactctgag ggtagccata atgaagaggc caaagaaccc 1200
gataagacaa ataagaaaga aggagagaaa acagatagag taggagttga agctagtgag 1260
gaaactcccc agacctcatc atcatctgct agaccaggta ctccttcaga ccatcagtct 1320
caggaagcat cacagtttga gaggaaagat gaaccaaaaa ctgagcaaac ggaaaaagct 1380
gaagaggaga ctcggatgga aaatagtcta ccagccaaag tgcccagcag aggggatgaa 1440
atggttgctg atgtccagca gcccctgtcg cagattcctt cagatacagc ctctcctctt 1500
ctcatacttc cacctcctgt tcccaatcct agtcctactc tccggccagt tgaaacacca 1560
gttgtaggtg ctcctggtat gggcagtgtt tccacagaac ctgatgatga agaaggtctc 1620
aaacatttgg agcagcaagc tgagaaaatg gtggcttatc tccaagacag tgcactagat 1680
gatgaaagat tggcatcaaa actgcaagag cacagagcta aaggagtgtc gattccattg 1740
atgcatgaag caatgcagaa gtggtattac aaagatcctc agggagaaat tcaaggtccc 1800
ttcaataatc aggagatggc agaatggttt caggcgggct attttactat gtctttattg 1860
gtgaagagag cgtgtgatga aagcttccaa cctcttggcg atatcatgaa aatgtgggga 1920
agggttccct tttctccagg tccagctccc cctcctcata tgggagagct ggaccaggaa 1980
cgactgacca ggcagcaaga actcacagcc ttataccaga tgcagcacct gcagtaccag 2040
cagtttttaa tacaacaaca atatgcacag gttttggccc aacagcagaa agcagcactg 2100
tcttcccagc agcagcagca gttggcactt cttcttcaac agtttcagac cttgaagatg 2160
agaatatctg atcagaacat cattccctca gtaactaggt ctgtgtccgt gccagatact 2220
ggctctatct gggagcttca gccaacagct tcacagccta cagtttggga aggtggtagt 2280
gtatgggatc ttcctctgga caccacgaca ccaggccctg ccctggaaca gcttcagcag 2340
ctagagaagg ccaaagctgc aaagctagag caagagagaa gagaggcaga aatgagggca 2400
aaacgggaag aggaagagcg aaagaggcag gaagaactcc gaagacaaca ggaggaaatt 2460

cttcggcgac agcaggaaga agaaaggaaa aggcgagagg aagaagaact tgcccgaagg 2520
aaacaggaag aggctctgcg tcgccagcgg gagcaagaaa ttgcattaag gcgacagcga 2580
gaagaggaag aaagacagca gcaagaagaa gctcttagaa gactggaaga gaggagaaga 2640
gaagaggaag aaaggcggaa gcaggaagaa ttgttacgca aacaggaaga ggaggctgca 2700
aaatgggccc gggaagaaga agaagcccag cgtcgattag aggagaaccg gctgcggatg 2760
gaagaggagg cagccagact ccggcatgag gaagaagaac ggaagagaaa ggagctggag 2820
gtccagcggc agaaggagtt aatgcgccag aggcagcagc agcaagaggc tctccggagg 2880
ttgcagcagc agcagcagca acaacagctg gcgcagatga agcttccttc ttcttcaacg 2940
tggggccagc agtccaatac aacagcatgt cagtcccagg ccacgctgtc gttggctgaa 3000
atccaaaaac tagaggaaga acgagaacgg cagcttcgag aagagcaaag gcgccagcag 3060
agggagttga tgaaagctct tcagcagcag cagcaacagc aacagcagaa actctcaggt 3120
tgggggaatg tcagcaaacc ttcaggtacc acgaaatctc ttctggagat ccagcaggaa 3180
gaggccaggc aaatgcaaaa gcagcagcag cagcagcagc aacaccagca accaaacaga 3240
gctcgtaaca atacgcattc caacctgcac accagcattg ggaattctgt ttggggctct 3300
ataaatactg gtcctcctaa ccagtgggca tctgacctag tcagtagtat ttggagtaat 3360
gctgacacta aaaactccaa catgggattc tgggatgatg cagtgaaaga ggtgggacct 3420
aggaattcaa caaataaaaa taaaaacaac gccagtctca gtaaatctgt aggtgtgtct 3480
aaccggcaga ataagaaagt agaagaagaa gaaaagttgc tgaagctctt tcagggagta 3540
aataaagccc aagatggatt tacgcagtgg tgtgaacaga tgcttcatgc ccttaatacg 3600
gcaaataact tggatgttcc cacatttgtt tctttcctga aagaagtaga atctccttat 3660
gaggtccatg attatatcag ggcctattta ggagatactt ctgaggccaa ggagtttgcc 3720
aagcagttcc ttgagcgccg tgccaaacag aaagccaacc agcagcgtca gcagcagcag 3780
ctgccacagc agcagcagca gcagccgcca cagcagccgc cacagcagcc acaacagcag 3840
gactctgtgt gggggatgaa ccacagtaca ctccattcag tatttcagac caatcaaagc 3900
aacaaccaac aatccaattt tgaggctgtg cagagtggca agaagaagaa aaagcagaag 3960
atggtccgag cagatcccag tttattagga ttttcagtca atgcatcatc ggagcgactc 4020
aacatgggtg aaatcgagac gttggatgac tactgagcac ctgccagtgg actggccatc 4080
cctctcctgt ctgccgacta tggagtctcc acctttggac acaacactta ctcaccattt 4140
actctttatc actctgcaac aaatcacaga accgatcatc tcaggctttt tcttctggcc 4200
ctttgtgtcc aagattcttt aatccatttt tgttggtgaa catctcagac tatagataag 4260
tggactggac cctgtgtctt gggggtggca gttgggatta ctccccaaca aggctgattt 4320
taggcagcat gtgttcactg tgctgtgatt tcatctactg tctcccagaa agtgtgttgg 4380
gatcggccat tagcagcttg ctttctcttg tcacttttt tcttctattt tgttttttct 4440
tcttcttttt ccccccatca gggcaaatgg tctaactggt gcaatcatga agagagttaa 4500
tggttaacag acattggcca ataacaaaac accccatgga ctgtgactcg agtatccaac 4560
aggcagtcag agctctcccg gtctgaaagt tgcattgcca ctgctaactt tgggattgca 4620
tcagagaggc cctgagtggg gttgagatga ggttggtttg gtttgatgtt acacactcct 4680
cacctgttct ttctgagtgt cctttctctg aaaggattta tgtttttctt cgttagatag 4740
tgacttctga gcaagctgat ctcccctggc atgctccaac ctgattggac aaaggaagct 4800
ctatggcctg ggagagagac tattcttaat ttttctttct tacaaaaact gatttttccc 4860
ataaatattt ttacttcaga ggactaggac cattttgttt tgggcccttc tgctgaaaat 4920
ttgtctcgtt taagaggcag ctagaatctt taccatatgt atgaatttgt ataatttcat 4980
ttttggatag ggataaactt ttgcttctga taaaagcctg gaatttcatc tggtcctcag 5040
agcattgcgt gtgtgtcttg ctgtagcccg gaaaaggttt tgtgtaaaga ttctgggatg 5100
gcaagttgtt tgccttttct gaaaagagaa catacagaac ctgtccatct ttaagacctt 5160
catccatgga atctactata caggaggatg cagtgggctg gaggggatgg gcgaaaatgg 5220
gagcaggaag cctggcctgg cttctggtca tggcctccta aaaccttaaa cttcaagtag 5280
aaatgtactc aagccctatt tataaacaaa tacttttcct gcctccacca aacccctaca 5340
gaacatcacc tggaattgcc actcacactg ggttggagtc attgggcagc tgtgcctgtg 5400
cgagaggtgc tgtggtctgg gcagcccctg gaaaagcacc tttgctgcct gtcattgttg 5460
cctgaagaag gctggagttg ctctgagagc agtttgggtt tggagtatta tatttggctt 5520
ctatttttat tattttggat caccattctc cctatccctt cttgcctccc tcccttctaa 5580
acatgtgtaa taactataca gagactgcta caaaattgta tatagttttt ggatcaaata 5640
gcatgagggg agaggaaacc attaaaaatt ggggctccta ctctcctttg ctttgtaaat 5700
tcaaaagttg ggggtgggta agagggatag ttaaaatgtt tacaaaactt taggctccct 5760
cggaactttt gccagtgtgg aggaaaataa aaagaacttt aaataaaatc tgattgtatt 5820
ctatctgagt gcacctcttg tactcacctt tatggaggct gagttctgca ctaaactgtt 5880

cctcttggta ccatggaaaa gctccaagca cccaagacat ggaggcagcc atggcttctt 5940
tctctgccag accacgtagc actggctggt tctgtatttg agaatgtaga ggtcaaggca 6000
gatgtcggaa tgctgatacc tgattctagt tatggatatg gaggcaagga gaagttgtta 6060
ggttaaccag cgcagtcctc cgtgcgtccc gaccgccgct gccctcactc ccggccaaga 6120
tggcatcctg tctggcac 6138

<210> 58
<211> 3190
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 1399169

<400> 58
cttcggggcc aagatggcga cgggaacggg caaacacaag ctgctaagca ctggccccac 60
agagccatgg tccatccgag agaagctatg tttagcatct tctgtcatga gaagtggcga 120
tcaaaattgg gtatcagtta gcagagcaat caagcccttt gcagaacctg gccgccctcc 180
agactggttc tctcaaaaac attgtgcttc ccagtactcg gagcttttag agaccactga 240
gacaccaaaa cggaaacgag gtgaaaaggg agaagtggtg gaaactgttg aagatgttat 300
tgttcggaaa ttgactgctg agcgagttga agaactaaag aaagtgataa aggaaaccca 360
ggagagatat agacggctaa agagagatgc agaactaatt caagctggac acatggacag 420
cagactggat gagctttgca atgacattgc aacgaaaaag aaattggaag aagaggaggc 480
tgaagtaaag aggaaggcta cagatgctgc ataccaggct cgtcaagcag taaaaacacc 540
cccccggagg ttacccactg tgatggttcg ctctcctata gattctgcct ccccaggagg 600
tgattatcca cttggggact tgactccaac cactatggaa gaggctacct ctggggtaac 660
ccccgggact ttgccgagta ccccagtcac ctcgtttcct gggattcctg acacccttcc 720
tccaggctct gcacccttag aagcccccat gaccccagta acagatgatt caccccagaa 780
aaagatgctt ggacagaaag caactccacc cccctccct ctgctgtcag agctcttgaa 840
gaagggcagc ctcctgccta ctagccccag actggtcaat gagagtgaaa tggctgtggc 900
ttctggccac ctgaacagta caggtgtcct cctggaggta ggcggggtcc ttcccatgat 960
acatggtggg gagatacagc aaacacccaa tactgttgca gcctcccctg ctgcatcagg 1020
tgctcccact ctttccggc ttttagaagc tggtcctaca cagttcacca cacctcttgc 1080
ttccttcact actgttgcca gtgagcctcc agttaaactt gtgccacccc ctgtagagtc 1140
tgtgtcccaa gctaccattg tcatgatgcc tgcgctgcca gcaccatcct ctgctccggc 1200
tgtctccact actgaaagtg tagctccagt gagtcaaccc gacaactgtg ttcccatgga 1260
ggctgtgggg gatccacata ctgtgactgt ttccatggac agcagtgaaa tatccatgat 1320
catcaattct atcaaagaag agtgttttcg atcaggggta gcagaggctc ctgttggatc 1380
aaaggctccc agcatagatg ggaaggaaga attagatctg gctgagaaga tggatattgc 1440
tgtgtcttac acaggtgaag agctggattt tgagactgtt ggagacatca ttgccatcat 1500
tgaggacaag gtagatgatc atcctgaagt gctggatgtg gcagcagtgg aagcagcact 1560
gtcattttgt gaagaaaatg atgatcctca gtccctgcct ggcccctggg agcatcctat 1620
ccagcaggag cgggacaagc cagtacctct ccctgcacca gaaatgacgg tcaagcaaga 1680
gagactggac tttgaggaaa cggaaaacaa gggaatacat gaactggtgg acatcaggga 1740
gcccagtgca gagatcaagg tggaacctgc agaaccagag ccagtcattt caggagccga 1800
aatagtagct ggagttgttc cagccacaag tatggagcca ccagaactca ggagtcagga 1860
cttagatgag gaactgggaa gtactgcagc tggagagatt gttgaagcag atgttgccat 1920
tgggaaaggc gatgagactc cacttacaaa tgtgaagaca gaggcatccc ctgaaagcat 1980
gttgtctcca tcacatggct caaatcccat tgaagatcct ttagaggcag agactcagca 2040
caagtttgaa atgtcagact cattgaaaga agaatcaggg actatttttg gaagccagat 2100
aaaggatgcc ccaggtgagg atgaggagga agatggtgtc agtgaagcgg ccagcctaga 2160
ggagcctaag gaagaggatc aaggagaagg ctacttgtca gaaatggata atgaacctcc 2220
tgtgagcgag agtgatgatg gcttcagcat acacaatgct acactgcagt cacacacact 2280
ggcagactcc atccccagca gccctgcttc ttcacagttc tctgtctgta gtgaggatca 2340
ggaagctatt caggcacaga aaatttggaa gaaagccatc atgcttgtat ggagagctgc 2400

agctaatcat aggtatgcca atgtcttcct gcagcctgtt acagatgaca tagcacctgg 2460
ctaccacagc attgtgcaga ggcctatgga tttgtcaact attaagaaaa acatagaaaa 2520
tggactgatc cgaagcacag ctgaatttca gcgtgacatt atgctgatgt ttcagaatgc 2580
tgtaatgtac aatagctcag accatgatgt ctatcacatg gcagtggaga tgcagcgaga 2640
tgtcttggaa cagatccagc aattcttggc cacgcagttg attatgcaaa catccgagtc 2700
tgggatcagt gctaaaagtc ttcgagggag agattctacc cgcaaacagg atgcttcaga 2760
gaaggacagt gtcccaatgg gctctcctgc cttccttctc tctctctttg atggaggaac 2820
caggggacgc cgctgtgcca ttgaagcaga tatgaagatg aaaaagtgaa gcctcagagt 2880
taccctcttt gagccgaacc taaaataaaa gtaaacaaga tagagcttgg gcttgcgggc 2940
ccagttccag aggtggaagt tacagaagag gaggtacctg ggccacacga catgagctgg 3000
aaaatctctc ttagagagtt ggagtagcac aattgcctgt tttagggcag aaaccatggg 3060
ctatgttaat gtcctaatgt gtagctagca gatcgtagct agtttgtatt gtcttgtcaa 3120
ttgtacagac tttttaaaaa aaacaaccac cagtgaaatg tgtgtgtata caataaactg 3180
aaaaaaaaaa 3190

<210> 59
<211> 1391
<212> DNA
<213> Homo sapiens

<220>
<221> unsure
<222> 1203, 1204
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte Clone No.: 1442069

<400> 59
ttttttttt ccagaggaga gtacaggtcg tgctgcagtt agttcattga aaactcattt 60
gctcttggag cagtcaggca gtgactgcct tcggcttttt ttctgctgac taagatctcc 120
tatagagagc tacaacaatg cccaaaagaa aggctgcagg tcaaggtgat atgaggcagg 180
agccaaagag aagatctgcc aggttgtctg ctatgcttgt gccagttaca ccagaggtga 240
agcctaaaag aacatcaagt tcaaggaaaa tgaagacaaa aagtgatatg atggaagaaa 300
acatagatac aagtgcccaa gcagttgctg aaaccaagca agaagcagtt gttgaagaag 360
actacaatga aaatgctaaa aatggagaag ccaaaattac agaggcacca gcttctgaaa 420
aagaaattgt ggaagtaaaa gaagaaaata ttgaagatgc cacagaaaag ggaggagaaa 480
agaaagaagc agtggcagca gaagtaaaaa atgaagaaga agatcagaaa gaagatgaag 540
aagatcaaaa cgaagagaaa ggggaagctg gaaaagaaga caaagatgaa aaaggggaag 600
aagatggaaa agaggataaa aatggaaatg agaaaggaga agatgcaaaa gagaaagaag 660
atggaaaaaa aggtgaagac ggaaaaggaa atggagaaga tggaaaagag aaaggagaag 720
atgaaaaaga ggaagaagac agaaaagaaa caggagatgg aaaagagaat gaagatggaa 780
aagagaaggg agataaaaaa gaggggaaag atgtaaaagt caaagaagat gaaaaagaga 840
gagaagatgg aaaagaagat gaaggtggaa atgaggaaga agctggaaaa gagaaagaag 900
atttaaaaga agaggaagaa ggaaaagagg aagatgagat caaagaagat gatggaaaaa 960
aagaggagcc acagagtatt gtttaaaact gccctatgta gtttcataat ttggtaacat 1020
gtaccttcat gttgtaaagt taatagagat aaatattttt atcaaaaatt ttataaacac 1080
agcctttctt tagcattgat ttaatttcag aacatcttca tattgattat tagccataaa 1140
gtttctaaca tgaaacattt atctataaat tttgtgatta tagtagtgga atacatagaa 1200
aannatatgc tttcaacttt gtgagtggat ttcgtgatgt gtagttatat gtcaatcttt 1260
ggttttaatt tactctttta tacatgtgat agttcataag tgagggatcc aaaacaaggt 1320
tcatccacat tcctgtctgc aggtgcttta taagaggtga ctatttcagt aatgtagggt 1380
aactatcttc c 1391

<210> 60
<211> 1125
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 1596668

<400> 60
caacaaagca aggaagacgg agtccgagcc tcgggggctc ctagcaacgg gccggggcgg 60
gagttccatg gagactgggg agcgcgcccg tctcatcctc atccttgtcc tccagcttct 120
ccttcgcatc cgacgcaacc ggcagcagcg ctgccgccgc gtcctcagcc accgctccct 180
cttcccacgg atgtgatctt cgtggtggaa agctaaattt taaaaccacc ccaatggatg 240
cagacagtga tgttgcattg gacattctaa ttacaaatgt agtctgtgtt tttagaacaa 300
gatgtcattt aaacttaagg aagattgctt tggaaggagc aaatgtaatt tataaacgtg 360
atgttggaaa agtattaatg aagcttagaa aacctagaat tacagctaca atttggtcct 420
caggaaaaat tatttgcact ggagcaacaa gtgaagaaga agctaaattt ggtgccagac 480
gcttagcccg tagtctgcag aaactaggtt ttcaggtaat atttacagat tttaaggttg 540
ttaacgttct ggcagtgtgt aacatgccat ttgaaatccg tttgccagaa ttcacaaaga 600
acaatagacc tcatgccagt tacgaacctg aacttcatcc tgctgtgtgc tatcggataa 660
aatctctaag agctacatta cagatttttt caacaggaag tatcacagta acagggccca 720
atgtaaaggc tgttgctact gctgtggaac agatttaccc atttgtgttt gaaagcagga 780
aagaaatttt ataattcacc acttaattgg ttagaatctc taactgagca ccttttaaac 840
ctgctgcaca ttggactcaa aaggaaaact ggaccaacaa taattgagga aatagactct 900
tttattcatt cacggctaca gtgtaagctc cagtcccttt ggattttatt ccaaaccttg 960
ctgtaatata aaaggaagtt tacaagacat gatattgctg cttttacaaa aggacattct 1020
atttattttc gcagtaattc tcatgtcccc ataagcagag ctgtcacagt gtgcactacc 1080
ttagattgtt ttattgtcgt cattgttatt tttttccatt ttgag 1125

<210> 61
<211> 3073
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 1977214

<400> 61
cggaattcgg ctcgagggcg actggagcgg ttccctcgca ggcggcgcca ttttgtgcta 60
ggagcctgat aaaaccggcc cggttctgtg gaaagtgggc ggcggagcca gggtccctgg 120
aatggcggag actctgtcag gcctaggtga ttctggagcg gcgggcgcgg cggctctgag 180
ctccgcctcg tcagagaccg ggacgcggcg cctcagcgac ctgcgagtga tcgatctgcg 240
ggcggagctg aggaaacgga atgtggactc gagcggcaac aagagcgttt tgatggagcg 300
gctgaagaag gcaattgaag atgaaggtgg taatcctgac gaaattgaaa ttacctccga 360
gggaaacaag aaaacatcaa agaggtctag caaagggcgc aaaccagaag aagagggtgt 420
ggaagataac gggctggagg aaaactctgg ggatggacag gaggatgttg agaccagtct 480
ggagaacttg caggacatcg acatcatgga tatcagtgtg ttggatgaag cagaaattga 540
taatggaagc gttgcagatt gtgtcgaaga cgatgatgct gataacctcc aggagtccct 600
gtcggatagt agagagctag tcgaggggga aatgaaagag cttccggagc agcttcagga 660
acatgctata gaggacaaag aaactataaa caatttagat acttcatcat ctgacttcac 720
tatattacag gaaattgaag agccatccct ggagccagaa aatgagaaaa tactcgacat 780
tttgggggaa acttgtaaat ctgagccagt aaaagaagaa agttccgagc tggagcagcc 840
atttgcacag gacacaagta gcgtggggcc agacagaaag cttgcggagg aagaggacct 900
atttgacagc gcccatccgg aagagggtga tttagatttg gccagcgagt caacagcaca 960

cgctcagtcg agcaaggcag acagcctgtt agcggtagtg aaaagggagc ccgcggagca 1020
gccaggcgat ggcgagagga cggactgtga gcctgtaggg ctagagccgg cagttgagca 1080
gagtagtgcg gcctccgagc tcgcggaggc ctctagcgag gagctcgcag aagcacccac 1140
ggaagcccca agcccagaag ccagagatag caaagaagac gggaggaagt ttgattttga 1200
cgcttgtaat gaagtccctc cggctcctaa agagtcctca accagtgagg gcgctgatca 1260
gaaaatgagt tctcccgaag atgactcgga tacaaaaagg ctttccaaag aggaaaaggg 1320
tcgcagcagt tgtggtagaa atttctgggt tagtggactc tcttctacaa ccagagctac 1380
agatttgaag aatcttttca gcaaatatgg gaaggtggtg ggcgccaagg ttgtgacaaa 1440
tgcccggagt cctggagctc gctgttacgg ttttgtcacg atgtccacag cagaagaggc 1500
cacaaaatgc attaaccacc tgcacaagac ggagctccac ggaaagatga tctccgtgga 1560
gaaagccaaa aatgaacctg tgggaaagaa aacctctgac aaaagagaca gtgacgggaa 1620
aaaggagaag tcgagcaaca gtgacagatc tacaaacctt aagagggatg ataaatgtga 1680
cagaaaagat gatgctaaga agggtgacga cggaagtgga gaaaagagta aggaccaaga 1740
tgatcagaaa cctggcccct cagagcgatc tcgagccaca aagtcaggaa gtcgagggac 1800
cgaacggact gtagtaatgg ataaatccaa aggggtgcct gtgattagtg taaaaacgtc 1860
cgggtccaaa gagagagctt ccaaaagcca ggatcgcaaa tcagccagca gagagaagcg 1920
gtccgtcgtg tcctttgata aggtcaagga gcctcggaag tcaagagact cagagtccca 1980
tagcagggtg cgtgaacgca gtgaacgcga acaacgcatg caggcgcagt gggagcgcga 2040
ggagcgtgag cggctggaga ttgcccgaga gaggctggcc ttccagcgcc agcggctgga 2100
gcgggagcgc atggagcggg aacggctgga gcgcgaacgc atgcacgtgg agcacgacgg 2160
caggcgcgag caggagcgca tccaccgtga gcgcgaggag ctgaggcgcc agcaggaact 2220
gcgctatgag caggagcggc ggcccgcggt gcggcggccc tacgacctgg accggcgaga 2280
tgatgcctat tggccggaag ccaagcgggc cgccctggat gagcgctacc attctgactt 2340
taaccgccag gaccgcttcc acgactttga ccacagggac cgcggccgct accccgacca 2400
ctcggtggac aggagagaag gttcaaggtc aatgatggga gaacgagaag gacagcatta 2460
cccagaacgc catggaggac cagagcgcca cggcggggcc tcccgcgatg gctgggggg 2520
ctatggctct gacaagagga tgagcgaggg ccgggggctg cctcctcccc ccaggggcag 2580
acgtgactgg ggggaccatg gccgaagaga ggatgaccgg tcatggcagg gcacggccga 2640
cgggggcatg atggacaggg atcacaagag gtggcaaggt ggcgagagaa gcatgtccgg 2700
tcactccggg cctggccaca tgatgaaccg aggaggaatg tcagggcgcg gcagctttgc 2760
cccaggcggg gcctcccggg gccaccccat cccacacggt ggcatgcagg gcgggtttgg 2820
aggccagagc cgggggagca ggcccagcga tgcccgcttc actcgccgct actgagtact 2880
tggaatcctg tgtcctgtct cgtggcaaca aggctatgtt ctgttaggag ttaccttaaa 2940
ctgtgtaaaa atatttttt ttaatctgct gccatattgt agctcaatac aatgtgaatt 3000
tgtttttcgt tttggggttt ttttttttg taataaatgt gtttccgttc acataccctt 3060
taaaaaaaaa aaa 3073

<210> 62
<211> 1446
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte Clone No.: 2181282

<400> 62
agagtaaaaa atgatcccca agaattcaga gaccaaaagc tggggaccct gaaaaaatac 60
cgtagcatta tgcccaaacc tatcatggtc atacccactt tggcctccct ggcttctcca 120
actacactac agtcccagat gcttgggggc ctaggacagg atgttttgtt aaataattca 180
ctcactccta aatatcttgg ctgtaagcaa gacaacagct cttcccctaa gcccagctcc 240
gtgttcagaa atggattctc tggcattaag aagccttggc acagatgtca cgtctgcaac 300
caccacttcc agttcaaaca gcaccttcga gaccacatga atacacacac caacagacgc 360
ccttacagtt gtcggatttg tcgcaagtcc tatgtacgtc ctggcagcct gagcacacac 420
atgaaacttc atcatggtga gaaccgtctg aagaaactca tgtgttgtga gttttgtgca 480
aaagtgtttg gccacatccg agtctatttt ggccatctga aagaagtgca tagggttgtg 540

atcagcactg agcctgcgcc cagtgaactg cagccaggag acataccaaa gaacagagac 600
atgagtgtgc gaggcatgga gggatcattg gagagggaaa acaagtccaa cctggaagaa 660
gacttccttc taaaccaggc agacgaagtc aaattacaaa tcaaatgtgg tcgttgtcag 720
attactgctc agtcttttgc ggaaataaaa tttcatttac ttgatgttca tggagaggaa 780
attgagggca ggctacaaga agggaccttc ccaggaagca aggggactca ggaagagttg 840
gtgcagcacg ctagccccga ctggaaaagg catcctgaga gagggaagcc ggagaaggtt 900
cattcctcct ccgaggaatc acatgcatgt ccaagactga aaaggcagct ccaccttcat 960
cagaatggcg tggaaatgct catggaaaat gaaggacccc agtcaggaac caacaagcca 1020
agggaaacct gccagggccc tgagtgtcct ggcctccaca cgtttctctt gtggtcccat 1080
tcaggcttta actgcctgct ttgtgcagag atgctgggac ggaaagagga cctcctccac 1140
cactggaagc accagcataa ctgtgaggac ccttccaaac tgtgggctat tttaaatacg 1200
gtctccaacc agggagtgat cgaactttcc agtgaagctg agaaatgaga ccccaaggca 1260
gcctggggtt aaggagagag ctctgccgcc accttccttc agagcttcgt gctttatggt 1320
ggtgcttagt cacaaagatc aaacaacagg attggtgtga gtgaacggaa atgatttttg 1380
tacatggttt tattttctta acgaaataaa atataagctc tcgaagcata tttttctaaa 1440
aaaaaa 1446